DOCUMENT RESUME

ED 154 596 EC 110 488

AUTHOR	Joiner, Lee Marvin; And Others
TITLE	A Technical Analysis of the Variation in Screening Instruments and Programs in New York State.
INSTITUTION	City Univ. of New York, N.Y. Center for Advanced Study in Education.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE	77
CONTRACT	OEC-0-74-7895
NOTE	108p.
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Diagnostic Tests; Elementary Secondary Education; *Evaluation Methods; *Handicapped Children; Identification; *Screening Tests; State Programs; Tests
IDENTIFIERS	New York

ABSTRACT

The monograph presents an analysis of screening programs for handicapped children in New York State school districts. An introductory section considers such factors as the general parameters of a screening program and characteristics of a good screening program. Major policy options for screening are seen to include decisions regarding universal or selective screening, use of screening tests of local origin, and examiners' competencies. Taxonomies for screening learning impediments are reviewed, along with a discussion of the variability of screening procedures and a listing and evaluation of instruments used in New York State. Suggestions are made for developing a public awareness program to be used in conjunction with the formal screening program. The final section provides overviews of 25 screening tests used in the state. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

The Graduate School and University Center
of the City University of New York

Center for Advanced Study in Education Teaching Resource Center
Graduate Center, 33 West 42 Street, New York, N.Y. 10036

ED154596

A TECHNICAL ANALYSIS OF THE VARIATION IN SCREENING INSTRUMENTS AND PROGRAMS IN NEW YORK STATE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Martin A. Hayott
Director

Developed under contract no. O.E.C.-0-74-7895 with the Bureau of Education for the Handicapped, U.S.O.E./D.H.E.W. awarded to the City University of New York, Graduate School and University Center, Case 06-77. The opinions expressed herein do not necessarily reflect the position or policy of the U.S.O.E./D.H.E.W. and no endorsement by that office should be inferred.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Martin A. Hayott

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

EC110480

Funded under contract with the Bureau of Education for the Handicapped, USOE, DHEW

A TECHNICAL ANALYSIS OF THE VARIATION IN SCREENING INSTRUMENTS AND PROGRAMS IN NEW YORK STATE

Developed by

Lee Marvin Joiner, Ph.D

Edited by

Emily Butler Kironde, Coordinator
Rebecca G. Graves, Research Assistant

CONTENTS

INTRODUCTION	i
Part I Introduction To Screening	1
Part II Major Policy Options For Screening	10
Part III Structures Or Taxonomies For Screening Learning Impediments	23
Part IV Variability Of Screening Procedures Within New York State	30
Part V Published Instruments Used In New York State Screening	35
Part VI The Combined Public Awareness And Formal Screening Option	43
Part VII Overviews Of New York State Screening Instruments	48
Bibliography	95

INTRODUCTION

Public Law 94-142 mandates a free and appropriate public education for all handicapped children, and provides that "all children who are handicapped will be identified, located and evaluated." In addition the law determines priorities which specify two prime categories for funding: "First, with respect to handicapped children who are not receiving an education and second, with respect to handicapped children with the most severe handicaps."

Implicit in this first priority is the need for efficient and effective means of locating, identifying, and screening those children who may have a handicapping condition. The New York State Legislature affirmed in Section 4408 of N.Y. State Education Law that screening is a vital first step toward providing adequate education for all the children in the state, and in 1975 directed the New York State Education Department to survey the state's screening procedures and practices.

The survey was conducted via a mail questionnaire sent to all school districts in the state. The resultant findings were compiled in a report to the New York State Legislature by Commissioner of Education Ewald B. Nyquist, entitled, A Survey Of Educational Screening Programs In New York State School Districts And Boards Of Cooperative Educational Services.

The New York Regional Resource Center has expanded this first effort by engaging in the process of examining, analyzing and evaluating the instruments and methodologies utilized by New York State school districts. The New York Regional Resource Center (NYRRC) anticipates that by providing this analysis of the variation in instrumentation

used for differential screening of handicapped children in New York State, legislators, as well as other state and local policy makers, will have a broader base of options from which to formulate recommendations and institute policies and procedures on screening.

The development of this monograph is in keeping with the New York Regional Resource Center's mandate to aid state program development through stimulating or developing intra-state and state capacity for child appraisal and educational programming. Screening is an assessment activity which should be carried out in a manner which 1) provides optimum cost-efficiency; 2) results in the identification of sufficient numbers of potentially handicapped children; and 3) provides information relevant for subsequent education evaluation and planning.

This monograph is an in-depth analysis of the New York State screening programs with respect to: the variation in instrumentation used for differential screening, the technical adequacies of the major published instruments used for differential screening, a comparison of appropriate usage of instruments with their actual application in school districts, and inter-comparisons between instruments use for screening in each category of handicap. A critical comparison is made of the New York State screening activities in relation to an evaluative standard based on a review and summary of the most recent professional recommendations concerning screening procedures.

In addition, this NYRRC publication outlines the implications of the variation in screening procedures and formulates policy recommendations and options for statewide screening efforts.

It is hoped that utilization of these findings and options will culminate in a reassessment of screening practices and will provide impetus for continuation of New York State's commitment to screening of young children.

PART I

INTRODUCTION TO SCREENING

GENERAL PARAMETERS OF A SCREENING PROGRAM

A. What is a screening program for learning impediments?

1. What are the purposes and presumed benefits of screening programs? Substantial professional confusion exists regarding the use of terminology describing various aspects of the process of locating handicapped children. In summarizing the Boston Conference of the President's Committee on Mental Retardation on Screening and Assessment of Young Children at Developmental Risk,

> ...I found no unanimity concerning the definitions of 'screening,' 'identification' and 'assessment,' nor pertaining to our concepts on the relationship of these three processes. Nor did I find agreement on the target population. (p.65)

It is no wonder, then if professional confusion exists about terminology and concepts, that the layman who is in a policy-making position is unclear on process and concept. Although complete consensus is lacking, we can specify the definition of some critical terms in a logically consistent fashion and ensure some clear communication. Hill (1970) defines proper identification procedures (hereafter referred to as screening) as:

> ....the use of a systematic procedure to identify provisionally those children from a population who manifest an attribute which is judged to require special attention. Screening involves the economical and effective testing of a population of children but it is not mass diagnosis. Screening programs are designed to identify those children for whom a more complete assessment is warranted.

Screening, then, is a first step in the process of locating handicapped

children. A variety of procedures including observation, checklist, and testing may be employed. As a consquence of screening, a number of children will be designated for further diagnostic evaluation. Those children who are in need of specialized services can then be located. While it is theoretically possible to conduct diagnostic evaluations on all children, and thus locate those in need of specialized services, such an intense approach is beyond the financial and professional staff resources of any school district, regional program, state or nation. In brief, screening procedures should focus on the identification of children who may need specialized services.

It should be kept in mind that screening procedures have potential utility for only mildly handicapped children. Severely handicapped children, i.e. the deaf, blind, serverely to profoundly retarded, or psychotic child will call sufficient attention to himself so that screening procedures are usually unnecessary, given the presence of a public awareness program. It is sufficient that primary service providers, in this case regular classroom teachers, have appropriate knowledge of and access to appropriate referral sources for diagnostic and specialized educational services. It is to the more mildly impaired child that screening programs are addressed.

2. Legislation on screening. An examination of Public Law 94-142 does not reveal any specific requirement that screening procedures be implemented <u>but</u> there is a clear requirement that handicapped children be located. Further, it would seem clear that the receipt of federal funding will sooner or later require that documentation of efforts to

locate children be provided. Some type of screening procedures seem clearly implied.

As will be shown in subsequent discussions of specific screening and diagnostic instruments, there may be some discontinuity between professional and legislative aspirations to locate and serve children with learning impediments and the actual "state of the professional art". For example, P.L. 94-142 asserts that diagnostic methods have advanced to a state of known effectiveness: "...developments in the training of teachers and in diagnostic and instructional procedures and methods have advanced to the point that given appropriate funding, state and local educational agencies can and will provide effective education and related services to meet the needs of handicapped children."

B. The sequence of events within and surrounding screening

A minimal series of requirements exists for the proper implementation of a screening program:

1. Those learning impediments for which children are to be screened should be identified.
2. Instruments of demonstrated effectiveness in screening for those impediments need to be selected and made available.
3. Staff need to be trained in the administration of those instruments.
4. Staff competent in the scoring and interpretation of the screening procedures need to be available in sufficient numbers to review screening results.
5. Professional diagnostic services need to be available in sufficient quantity to see children "referred" by screening procedures.
6. Remedial services need to be available for those children located who are determined through diagnostic evaluation to be in need of specialized services.

If all of the above conditions cannot be met no benefits will derive from screening procedures. Screening is one step (not a necessary step) in process of identifying children in need of specialized services. If any step in the process is nonexistent or unavailable in sufficient quantity, then there is no reasonable purpose for screening. Professional staff trained in the interpretations of screening results, diagnostic staff, and specialized programming staff are the service elements most likely to be overloaded when the implementation of a mass screening program in considered. Experience in medical as well as educational programming supports this point of view.

C. General characteristics of a good screening program

1. How extensive should a screening program be?

In considering this issue a number of concerns arise. The overriding issue with screening is cost-effectiveness. A screening program must result in the placement of children in specialized programs without prohibitive cost. A program in which 1,000 children were screened, 100 seen for diagnostic services, 10 found in need of services, and 1 child received specialized services would be hard to justify in terms of its cost. The entire programmatic cost should be assigned to the 1 child who received services and this would be an excessive financial burden.

The following general factors should be considered in the designation of screening procedures:

a. Age. Earliest detection and intervention in the case of learning impediments is desirable. If problems are detected early, the best chance would seem to exist for remediation and the cumulative effects of the impediment might be prevented. However, both human and technological considerations exist which must be considered. The younger

the child, the greater the likelihood that he will be "untestable" or present problems in the testing situation which will result in referral. Definitive diagnoses are more difficult to arrive at with younger children. Finally, specialized services tend to be least well established for younger children. Thus, while theoretical considerations might lead toward a concentration of screening effort to younger children, practical considerations may lead toward later implementation of screening procedures.

b. Frequency. A single time or age for screening is unlikely to be effective for all learning impediments. Some aspects of learning disabilities are essentially undetectable until a child is of second or third grade age. For example, hearing (auditory acuity) problems may develop at any age and periodic repetitions of screening are required. An appropriate screening program cannot all be implemented at a single age and some aspects of it will need to be conducted on a repetitive basis.

c. Proportion of Population. Incidence rates on the prevalence of many learning impediments are inaccurate or non-existent. Frequently screening programs are conducted on an entire intact population, i.e., kindergarten age children. It may, however, be more appropriate to screen samples of an intact population for purposes of: 1. Development of screening instruments, 2. Development and evaluation of diagnostic procedures, 3. Development of service programs to meet expected needs. Many of the aspects of the delivery of specialized educational services lack full development.

2. When and where should screening be conducted?

Efficiency requires that screening be conducted as closely in conjunction as possible with the child's regular classroom education. The child's regular classroom teacher would seem to be an excellent screening resource. She can be trained to administer group tests and fill out checklists. Both procedures have excellent potential for high cost-effectiveness. In terms of planning for the implementation of specialized educational services, it would be most desirable to identify and complete diagnostic evaluations in the spring of the year so that effective plans could be made for the extent and types of specialized services required

for the following year. However, this approach has the effect of overloading diagnostic capabilities at one time of the year or requiring the employment of many extra diagnosticians at one time of the year. Quality diagrostic services cannot be obtained on a part-time basis. To avoid overloading diagnostic resources, screening should be conducted on a continuous basis within any administrative unit.

3. What levels of staff skill should be required?

Since a major intent of screening procedures in cost-effectiveness the use of personnel with limited training is implied in the case of individually administered screening procedures. The use of paraprofessionals is controversial. However, pretraining on test administration, competency evaluation of testing skills, the observation of trained examiners, and the observation of the examiner in real life situations are necessary to insure accuracy in screening. Hearing acuity screening with the use of pure tone audiometer is conducted by professionals. The Denver Developmental Screening test for preschool age (0 to 5), was developed for use by paraprofessionals. Considerable controversy exists in the professional literature on the accuracy of Denver results when nonprofessionals administer the test.

4. What are considered reasonable and realistic time and cost factors?

Virtually no literature exists on cost and time factors involved in screening for learning impediments. In discussing mass screening of the zero to five populaiton, including medical and laboratory costs, Tarjan (1973) estimated a cost of $100 per child. This cost is based on screening three million babies per year. Tarjan argues that this cost level is justifiable.

5. What is the general state of the art regarding screening instruments?

   a. Basic technical requirements have been established for evaluating screening devices.

   b. No single instrument will screen for all types of learning impediments.

   c. Socio-cultural problems exist with most screening instruments.

A good screening instrument should meet the following requirements:

   a. It should require relatively little time to administer. A diagnostic evaluation is measured in hours and days. Screening should be measured in minutes.

   b. It should be demonstrated that persons with limited training can administer the test with accuracy.

   c. The materials required should be inexpensive.

   d. The instrument must be reliable.

   e. The instrument should refer for further diagnostic evaluation the major portion of those children who are in need of specialized programming. Further, the numbers of children referred must not overtax diagnostic services so that cost-effectiveness is lost.

The development of screening instruments is in its infancy. It is fair to state that there is no single screening instrument that meets all of the requirements set forth above for even a single impediment or even less so for a constellation of learning impediments. There are a number of reasons why this is true. Primarily, the cost of screening instrument development is prohibitively high. To develop an effective screening device would require taking the following steps:

- Develop and conceptualize an area for screening.
- Develop items for the screening instrument.
- Train staff to administer and interpret the instrument.
- Establish reliability.

- Administer the screening device to a large number of children in order to identify sufficiently large pools of children with various impediments.

- Conduct diagnostic evaluations on all children screened to determine the rate of false positives and misses.

Additional complexities exist. Many different conditions have effects which can appear similar in both screening and diagnostics. Mild mental retardation, emotional disturbance, specific learning disabilities can all produce children with similar patterns of test result. Diagnostic errors are thus common. A child with a mild hearing impairment may appear to be mentally retarded, learning disabled, emotionally disturbed or show elements of all three conditions. Many mild hearing impairments are intermittent in nature. The child may pass a hearing screening test one day and fail it the next when he has fluid and infection in his ears. Finally, most approaches to child evaluation, including screening, are based on determining how well the child has learned from his environment. This carries the assumption that there are large similarities in environmental background. Most test are targeted toward middle environmental ground (i.e. middle SES). Children with socio-culturally different backgrounds sometimes have had little or no exposure to certain content items on these tests. It is also likely that learning impediments have different effects in different cultural circumstances. Hard evidence of these effects is lacking. With such monumental complexities it is no wonder that the development of screening instruments is in its infancy.

6. What are the arguments against a formal screening program?

It is somewhat difficult to oppose efforts to locate children with learning impediments. It would seem to be similar to opposing services

for handicapped children. However, some serious considerations might be reviewed:

- There is little evidence that early screening procedures for the identification of children with learning impediments have any effectiveness in assisting in the location of such children.

- There is little evidence to support the effectiveness of programs of specialized service for the child with mild learning impediments, and these are the children located in screening programs involving instruments or tests.

- There is danger of increasing public expectations with the initiation of massive screening programs. The public will tend to expect far more than the state of the art can deliver. The belief will develop that a child has no learning impediments if he passes screening. Such expectations may prove contrary to later reality.

- Minority and culturally different children may be perceived as victims of insensitive and inappropriate testing by their parents and other advocates. There have been no dramatic successes in eliminating cultural biases in tests because the scope and depth of cultural differences and their effects on learning and test performance are not fully understood. Therefore, nearly any examination item is open to challenge on these grounds.

PART II

MAJOR POLICY OPTIONS FOR SCREENING

Before one is able to arrive at a rational decision concerning specific requirements for a screening program, it is necessary to answer some basic policy questions which relate to the distribution of resources within an educational system. The following are basic policy questions:

1. Is it desirable to conduct universal screening for all children or should screening be a selective activity targeted to fewer than the total population of school children?

2. Will the screening battery be constant or variable? Will the same array of tests be used with each student or will tests be permitted to vary in purpose and content according to perceived needs of students?

3. Who will be directly responsible for the administration of the screening test? Will these persons be diagnosticians, classroom teachers or paraprofessionals?

4. When will screening be conducted, at pre-kindergarten, kindergarten, grade one, or at multiple levels?

A. Universal or Selective Screening?

This involves basic logistical and economic considerations. For even a minimal screening program if adopted on a universal basis involves the allocation of considerable fiscal resources. Some of this can be minimized by using paraprofessionals and volunteers to aid in administering tests. But tests must be acquired, assembled, and the results recorded, stored, and interpreted. Professional man-hours above and beyond those devoted to acquiring the observations

and data themselves are entailed in screening.

A very important consideration in deciding whether to adopt a policy of universal screening is the degree to which resources are available to follow-up on the screening results. In some parts of the United States, local education agencies have resisted intensified "child-find" effort because the ultimate responsibility for providing programming for children diagnosed as having learning impediments falls upon local school districts. Providing appropriate programs for children identified in universal screening and diagnostic follow-up places added financial burdens on local schools which previously have been operating without services for "unidentified" children. Also, diagnostic and resource staff may function under a condition of regulated case loads. If this condition prevails, screening efforts only result in longer waiting lists for diagnosis, unless additional resource and diagnostic staff is provided for thorough supplementary funding.

What is implied by the above considerations is that commitments to universal screening produces secondary expense factors which are not always considered in advance by policy-makers. Lack of follow-up diagnosis and programming for identified children can produce accelerated conflict within the system over the allocation of limited resources.

Selective screening, on the other hand, may or may not result from the presence of a formal policy of screening. Selective screening may occur

through default, a consequence of the absence of a comprehensive screening policy. There are, however, rational bases for defending a selective screening policy above and beyond the fact that it is less expensive than a universal program.

In selective screening, the parent, physician, social worker, or teacher becomes the primary screening instrument. Through evidence accumulated as a consequence of daily encounters with children, these persons may make a prior assessment of the likelihood that a child's performance is being hindered by some learning impediment. For children who are suspected of having unidentified learning impediments more formal procedures may be applied to begin isolating the problem.

The effectiveness of a selective screening system rests upon the extent to which pre-kindergarten and kindergarten teachers are willing and able to identify child behaviors that may be indicative of learning impediments. There is reason to believe, however, that teachers vary in competence to such a degree that one cannot be sure that their judgements are sound. Our greatest concern would be for those who are overlooked because of their passive temperaments and otherwise conforming behavior. These might include some children with developmental disabilities, such as mild mental retardation, withdrawal syndromes, mild hearing impairments, auditory perception disorders, and physical health problems. Of necessity, a selective screening program would require in-service courses for teachers on identifying behaviors associated with learning impediments and an intensified community awareness program.

Improving child observer skills connot overcome the basic disadvantage of a selective screening system which decreases the likelihood of early identification and early intervention. However, the successful results of early intervention are based on developmental theories rather than solid factual data as to the positive consequences of early intervention programs.

At the present time in New York State, the screening program could be characterized as selective. Only 24 percent of the schools reported operating screening programs involving formal assessment procedures at levels pre-kindergarten through grade one. Information on public awareness efforts, made as a part of the screening process, was not gathered on the New York State Survey.

B. Constant or Variable Screening Battery?

A second area for policy making is whether the screening battery should be constant within and between school districts. Were the state of New York to specify the exact tests that were to be applied in screening, it would be characterized as a constant testing policy. The question of whether to adopt a constant or variable testing policy is independent of whether or not screening is to be universal or selective. A constant battery of tests can be used with part of or the entire population of children.

Constant screening batteries have some appealing features. First, they allow for direct comparisons to be made between districts, schools, counties, and regions within the state. Because of consistencies in

data base, meaningful interpretations of variations in incidence of incipient learning impediments can be made. Such a base permits the compiling of statistical summaries and may meet certain reporting requirements of federal funding sources in an efficient way. Having a constant battery of screening tests also reduces confusion among professionals, enhances their ability to plan on a statewide basis, allows for steady improvements in staff skills because of their increasing familiarity with the instruments over a period of time, and permits the development of standardized in-service packages relating to the conduct of screening programs. A final important feature of constant screening programs is that they allow closure to occur on the question of whether or not a particular child has been fully screened. The response to this question is "yes" if the child has been recorded as having responded to all assessments in the battery. In general, a constant screening battery would be a closed-end process.

The basic dilemma posed by adopting a constant screening battery policy is that no single screening battery has yet been designed that will adequately detect the several learning impediments in a cost-efficient way. Comprehensive screening packages such as DIAL, which was developed in Illinois, have proved unsatisfactory. More adequate tests that are narrow in scope and targeted to specific problems, however, multiplies testing costs for time, material, and professional resources.

A variable screening program, in contrast, would be one which

permits each child to be exposed to a different number of tests or different kinds of tests depending on the educator's perception of local and individual needs. A wide open variable system would allow schools to use whatever test they deem appropriate for screening a particular child or group of children. The tests selected may be published or unpublished and of local origin.

A middle of the road option is to prohibit the use of unpublished tests of local origin because of their likely inadequacies. A discussion of this issue follows. In addition to this prohibition, an array of tests might be specified from which school districts could choose according to local and individual needs. This would allow for limited comparability between districts, although there would be an absence of complete data such as would occur under a constant screening policy. School staff would be faced with deciding which tests should be used in particular instances and the assumption would be made that they could arrive at this decision competently.

C. Screening Tests of Local Origin

There is little justification for encouraging school districts to continue the practice of designing their own screening instruments. First, public schools rarely retain staff who are skilled in psychometrics and developmental theory, both of which are crucial to designing adequate screening instruments. The New York State Survey data discloses the general level of testing expertise that prevails in the schools. There is evidence of an inability to even select appropriately among tests already published. Most of the published tests reportedly used for

screening in New York State were inappropriate for the stated purpose: i.e., tests were used to screen conditions for which they were not designed, nor capable of detecting. Only in the larger school systems might one expect to find personnel sufficiently prepared to develop adequate screening devices.

Test development requires studies of reliability, validity, and revisions of tests before they are disseminated. In local school test development such studies are rarely conducted. While lacking the training to address themselves adequately to the reliability of tests, schools lack the resources to design and conduct studies relevant to establishing the validity of an instrument. Establishing the validity of screening instruments also requires the use of longitudinal designs which follow children over a segment of their school carrers. The economics of education and the realities of the school calendar combine to mitigate against long-term test development work. Professionals typically need an instrument for immediate application. Staff turnover produces further complications for improving screening skills.

Rhetoric exists to the effect that tests should be developed locally in order to reflect local conditions and idiosyncracies. However, there is no evidence whatsoever to suggest that:

1) the nature of these local variations can be identified and analyzed for test development purposes, or

2) that locally developed tests addressed to local idiosyncracies are in any way superior to those available in the national marketplace.

It is unproductive to design local screening devices and arrive at local definitions of learning disabilities, mental retardaion and so forth, and to pick and choose sub-tests from among several tests in order to assemble a new test. What is achieved by this cutting and pasting is a test inferior to any single original test. This is assuming that no extensive reliability and validity analysis is done on the composite test. The reason for this dilemma is that sub-test, for example, "Copying" from the Metropolitan Readiness Test, have low reliabilities and validites. This reflects a basic principle of the test and the amount of content and behavior sampled. Subtest results rarely prove psychometrically adequate in any research that has been conducted on this question. The Metropolitan Readiness test is one of the few tests that issues a warning to this effect along with the administration instructions. Some tests, in contrast, promote the erroneous doctrine that each sub-test yields information which has independent merit even though published research on this exact subject disproves these assertions.

In summary, what are some of the consequences of encouraging or permitting schools to develop local screening batteries?

1. Very expensive in dollars and cents if good tests are to be produced.

2. Inefficient duplication of activities among districts, and many "rediscoveries of America" will emerge.

3. Very little comparability of results among districts will make it impossible to construct a meaningful profile of the early childhood education needs of the state.

4. Inefficiencies may occur and children with problems may remain undetected; or children may be incorrectly identified as needing diagnostic follow-up.

New York State utilizes variable screening test batteries. A heavy reliance on tests of local origin is documented by the State Education Department's Report to the New York State Legislature (1975). Also revealed in the report is a pervasive discontinuity among school districts in the tests that compose their screening batteries, and great variations even within single learning impediment categories, such as mental retardation (a total of 59 different screening tests were reportedly used in New York State to screen for mental retardation).

D. Competencies of Examiners

Nearly all screening procedures require that examiners receive some training in test administration, observation, and score recording. They must also be aware of standards for success and failure of a student so they can identify students who need to be referred for more detailed diagnosis. Most screening tests entail some subjective judgements by examiners and, if young children are involved, examiners must be sensitive to conditions which might impeded students in performing adequately. For instance, management problems associated with the structure of the DIAL Screening Program cause dissatifaction because children were difficult to control in the testing situation.

A generalization that tends to hold up in testing situations is that the more objective the test, the less the training requirements for the person administering the test. For instance, the Peabody Picture Vocabulary test requires only that a child point to a picture depicting the object or concept names. The examiner provides a verbal cue of the form: "Show me ...", records the response, turns the page of the test booklets and repeats the procedure. Tests of this type require little training to administer

and score. The Wechsler Intelligence Scale for Children, a test reportedly used for screening in some school districts in New York, is more complex and generally requires the skills of a certified school psychologist for administration and scoring.

The general trend has been to use screening instruments that can be administered and scored by the classroom teacher or even paraprofessionals without extensive training experiences. Therefore tests that entail limited subjectivity and that can be administered in group settings are sought. Once a decision has been made to use teachers or paraprofessionals as examiners, a number of screening instruments are automatically excluded from consideration for adoption as screening devices.

E. At What Grade Level to Screen?

A fourth major policy option concerns the appropriate level at which to screen children for learning impediments. For early childhood screening in New York State we find examples of pre-school, first grade, kindergarten and repeated screening in use.

Much of the current rhetoric in special education suggests that it is advisable to institute screening programs as early as possible in the life of the child, sometime during the pre-kindergarten years. The first dilemma induced by this position is that there is a general lack of reliable and valid instruments that have been proven effective with very young children. Most of the published tests used in New York State screening are unsuitable for pre-kindergarten use or lack psychometric information on very young children. Developing test procedures for pre-kindergarten children is also quite a formidable task because of the immaturity of the subject tested and the limited repertoire of competencies

that they can be expected to display. There exists a very narrow range of possible test items that are developmentally suited to this population. In addition, pre-kindergarten age children are noted for producing unreliable responses in test situations.

If a decision is made to screen pre-school populations, the logistic problems associated with launching such a program become monstrous. If parents are asked to bring their children to a central location for screening, the children who are most in need of early intervention and diagnosis more frequently fail to appear than children with low risk of learning impediments. Typically, the better educated, more affluent parents are more apt to bring their children to central screening programs.

If pre-school programs provide the population for screening we encounter a maldistribution similar to that noted above. Children who are most in need of early childhood compensatory program often fail to become enrolled in these programs. Therefore, again, we are in a position of screening a population of young children whose expected rate of occurrence of learning impediments is low. We thereby incur substantial costs without receiving maximum benefit of screening.

Outreach type programs which attempt to take the screening process into the home are extremely expensive unless a two-stage screening is conducted. In the first stage, parents are asked to respond to a limited number of simple questions concerning their child's physical status and behavior. On the basis of these data, a selective screening program is directed toward children who appear higher risks according to the parents response to preliminary survey questions. In order to be successful, the community must be canvassed in the first stage, much like a census.

In fact this process might be termed a "high risk child" census.

A final problem associated with pre-kindergarten screening is that a certain number of "false negatives" will emerge due to instrument unreliabilities and the additional fact that some problems only begin to become visible after the child has entered more formal school settings (e.g. mild mental retardation, school adjustment problems, social problems). A false negative is a case which has passed initial screening but later displays failure patterns associated with learning impediments. In response to this some schools screen at pre-kindergarten and again in kindergarten and/or first grade. In New York State, 28 of 736 school districts responding to the 1975 state survey reported screening only at pre-kindergarten; 31 of 736 at both pre-kindergarten and kindergarten; 6 at all levels, pre-kindergarten through grade one.

Repeated screenings result in more reliable results and are less likely to miss students who are in need of special compensatory services. However, repeated screenings elevate cost levels beyond what might be considered reasonable by some school districts.

The most common screening pattern in New York school districts is to screen only at kindergarten level; 83 of 736 school districts reporting have pursued this option. At the kindergarten level more published tests become available for application and the normative basis for making comparisons and interpreting performance becomes more adequate. The tendency among those schools targeting screening to the kindergarten-level child is to schedule screening activities near the close of the kindergarten year. An argument in favor of this scheduling is that kindergarten is designed for the purpose of enhancing instructional readiness skills and

some problems that might be noted at the beginning of the year will disappear as a consequence of the instructional program. However, for those children who turn up as high risk at the close of the kindergarten year, options will disappear as a consequence of the instructional programs. However, for those children who turn up as high risks at the close of the kindergarten year, options for modifying instructions or injecting special corrective procedures become reduced. Most school districts fail to support "transitional" programs, a type of further readiness experience between kindergarten and first grade situations with perhaps some resource service. Kindergarten is rarely repeated, but these children may be unprepared to begin more traditional instruction in first grade placements with any likelihood of success.

The least acceptable option is to screen children at grade levels above kindergarten. In New York State 11 of 736 school districts report screening children at grade one. Once the child has entered into the numerical grade levels, the onset of failure, maladaptation, and other problems often multiply rapidly. These early academic failures set the tone for later performance and begin the emergence of a somewhat inflexible pattern of school failure. Habits are established which may be maladaptive but very firmly entrenched nevertheless. Subsequent intervention must deal with extinguishing those behaviors before the child can benefit from remedial or compensatory instructions.

PART III

STRUCTURES OR TAXONOMIES FOR SCREENING LEARNING IMPEDIMENTS

A. Variations in Taxonomies of Learning Impediments

There are a number of organizing structures around which a screening program for learning impediments could be designed. This wide range of options is reflected in the diversity that exists in screening programs among states and among communities within states. Usually the organizing structure upon which a screening program is based represents some taxonomy of disorders which one seeks evidence of among children screened. Specific "functional deficits" such as limited short-term memory, broad syndromes such as mental retardation, and loose typologies such as "high risk" may provide the basis for screening. It should be recognized that much uncertainty regarding how to operationalize a screening program in resolved once some taxonomy of disorders is adopted as a structure around which testing is to be conducted.

Educators and psychologists have thus far failed to arrive at consensus concerning which organizing base should be used for characterizing "learning impediments." Because there is no single structure which has been shown to be most efficient or viable for organizing instruction, testing, research, and curriculum development, it follows that there is an element of arbitrariness in designing a screening program. To further complicate matters, education practices are constantly changing in response to new insights on how to best classify the problems that adversely affect learning and to various pressures which emerge from the political context within which education functions.

B. Categorical Taxonomies and Associate Criticisms

Traditionally, medical, psychiatric, psychological, sociological, educational, and psychometric criteria have been used to classify children as "disabled." These medical and psychiatric disability categories have tended, historically, to provide the structure for special education services. Even then there has been a lack of agreement among authorities concerning what to call and how to define these "disabilities." The five most widely used basic textbooks in the education of exceptional children disagree on the number, names and definition of various "disabilities."

A recent "streamlined" version of the traditional category structure used for programming is as follows:

1. Behavioral Disabilities - These include the emotionally disturbed, socially malajusted, and juvenile delinquent. Associated with this are "a variety of excessive, chronic, and deviant behaviors ranging from impulsive and aggressive to depressive and withdrawal acts which (1) violate the perceivers expectations of appropriateness, and (2) which the perceiver wishes to see stopped" (Graubard, 1973).

2. Specific Learning Disabilities - Various states use terms such as educational handicap, extreme learning problems, communication and intellective deviations, perceptually handicapped, etc. Public Law 91-320, "The Learning Disabilities Act of 1969" defines learning disability as... "children with special (specific) learning disabilities exhibit a disorder in one or more of the basic psychological processes involved in understanding or using spoken or written language. These may be manifested in disorders of listening, thinking, talking, reading, writing, spelling, or arithmetic. They include conditions which have been referred to as perceptual handicaps, brain dysfunctions, dyslexia, developmental aphasia, etc. They do not include learning problems which are due primarily to visual, hearing, or motor handicaps, or mental retardation, emotional disturbance, or environmental disadvantage..."

3. Visual Disabilities - Includes partially sighted and legally blind (legal definitions); partially sighted and braille readers (educational definitions).

4. Hearing Disabilities - Includes the deaf and the partially hearing.

5. Oral Communication Disabilities - Includes both language disorders (such as childhood aphasia) and speech disorders or impairments.

6. Crippling and Health Disabilities - Includes children with disorders of the central nervous system (e.g. cerebral palsy); children with musculoskeletal conditions (e.g. muscular dystrophy); children with cogenital malformations (e.g. malformations of the heart); children with other crippling and health conditions (e.g. asthma).

7. Mental Retardation - "Refers to significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior, and manifested during the developmental period" (American Association on Mental Deficiency).

The above system is an attempt at providing a rational structure which communicates something about the nature of the disorder and suggests similarities of educational needs among members of a particular category. During recent years, however, there has been a tendency to discredit "disability" categories and replace them with systems that are "deficit specific" or "functional." These systems direct attention toward the exact learning or behavioral deficts, such as "laterality confusion."

A number of weaknesses of the "disability" category system have been discussed at length in recent professional literature. These criticisms of the category system include:

1. Lacks Educational Relevance

2. Labels Children

3. Places a "Stigma" on Children

4. Fails to Consider Differences Among Children Within A Category

5. Promotes Illusion of Exaggerated Differences Between Categories

6. Produces "Self-Fulfilling Prophesies"

7. Contributes to the Social Isolation of Some Children

8. Lacks Egalitarianism

C. Functional Taxonomies as an Organizing Basis

Alternative systems for organizing learning and behavior problems have begun with an analysis of behaviors themselves rather than people. In addressing their concerns to behaviors, authorities have sought to avoid assumptions concerning "causes" of behaviors, for example, the idea that academic deficits are "caused" by mental retardation. They have sought to avoid such postulation because once a causal condition has been diagnostically verified, it still leaves the educator in the position of asking what to do next. Educational treatments have not been directly implied by causes.

The alternative or "functional" approach recognizes that an educational program does not "treat" mental retardation. What the educational program does instead is to attempt to remediate adaptive behavior deficits or various cognitive skills or aptitudes. Therefore, functional approaches to classification seek to determine which specific abilities and deficits prevail within each child. Theoretically, programming can then be organized around these fairly concrete entities. In this case, diagnosis becomes the basis for curriculum and grouping decisions.

Educators who identify with the "functional" approach are still left with the problem of arriving at a viable taxonomy. What this means is that there are an infinite number of specific "deficits" that can serve as focal points for diagnosis and instruction and these "deficits" can be organized in a great number of ways. No functional taxonomy has been

fully validated or accepted professionally. Some have received greater credibility and dissemination, such as Guilford's "Structure of Intellect." Others in the professional rhetoric and practice, however, have tended to comprehensiveness and have suffered from redundancy and conceptual glut. Many are only "quasi" empirical constructs, such as "mixed lateral dominance."

The following are some examples of concepts which could serve as an organizing structure for a functional approach to diagnosis and remediation (from Gardner, 1977):

1. Sensory and Perceptual Discrimination Deficits
2. Affective Behavior Deficits
3. Deficits in Age-Relevant Personal and Social Responsibility Behaviors
4. Laterality Problems (from the New York State Survey)
5. Language and Cognitive Behavior Deficits (from the New York State Survey)
6. Deficits in Play and/or Social Interaction Skills.

These are but a few illustrations of the many kinds of "learning impediments" that can provide a base for the functional approach to classification. Clearly this position would dictate a different approach to screening than if the traditional "disability" system prevailed.

Not only should a screening program reflect some decision as to the appropriate and best organizing structure in terms of our present state of knowledge, it is important to realize that an adequate screening program must reflect the structure of available intervention services. If the schools' compensatory and remedial programs are organized along traditional categorical lines, a screening program should be targeted to those conditions and make use of tests that are designed for the purpose of detecting them. Screening for conditions that cannot readily be related to existing or planned

intervention programs results only in the acquisition of descriptive, statistical data at a relatively high cost. Another way of stating this is that the screening program should not be more innovative than the programming structure.

D. The New York State Screening Structure

The importance of the above discussion becomes clear when we consider the twelve (12) "learning impairments" around which the New York State systems appears devoid of any rationale for including or excluding specific elements of either system.

For example, learning disabilities, mental retardation, brain-damage, emotional disturbance, and cultural disadvantage are identified as "categorical handicaps" consistent with section 4408 of the State of New York Education Law. Conspicuously absent from among these, however, are very important categories for traditional screening such as "visual disabilities, hearing disabilities, oral communication disabilities, and crippling and health disabilities. Brain damage, on the other hand, is rarely used today as a viable, independent, categorical handicap. It is typically subsumed under the more general category "learning disability" if it is a "minimal brain dysfunction," difficult to verify using present levels of sophistication in medical and psychological testing. But if the "brain damage" is major and clearly identified through medical diagnosis, the problems would be classified as a crippling and health disability.

While certain important categorical handicaps were excluded from the New York survey, "learning process components" of marginal educational relevance, according to the research literature, were included. Examples of marginally relevant "learning process components" would be laterality

and directionality. Although a vast and turgid literature exists concerning these constructs, they certainly cannot be considered of sufficient importance to be prime targets for screening programs. If they are to appear in educational assessment, they should appear at a later stage than screening i.e., in the more detailed diagnostic follow-up which must accompany any viable special education programming. Visual, auditory perception and sensory motor integration are also diagnostic constructs.

Previous academic experience and general behavior, on the other hand, could be a matter of record. Teachers are expected to report periodically on the progress of children in these areas. To screen for these phenomena would be redundant on activities that teachers already engage in. Certainly teachers should be aware of these problems since they are the ones who observe and record them. Similarly, previous academic experience should be a matter of record, and detecting evidence of problems here again involves only reviewing information routinely collected on children.

It is interesting that the categories visual and auditory acuity appear in conjunction with visual and auditory perception in the New York State Survey. Schools were asked to identify tests used to screen simultaneously for these problems. The fact that no tests were identified that are appropriately used for screening for visual and auditory acuity is not surprising since the present author is unaware of any devices that screen for both perception and acuity. For visual acuity, the traditional screening device is the Snellen chart, and for hearing acuity, the watch tick test or the whisper test. The fact that all school districts failed to mention either of these suggests that they provided responses of limited validity to the survey. Nearly all schools in the United States expose children to these tests routinely.

PART IV

VARIABILITY OF SCREENING PROCEDURES WITHIN NEW YORK STATE

The variability in screening procedures among school districts in New York State is extreme. According to the New York State survey 151 different instruments are used for screening, the majority of which were unpublished instruments of local origin. Moreover this extreme variability in instrumentation reflects the state of affairs that prevails in only 177 of 736 school districts surveyed; the remainder of the school districts indicated the complete absence of formal early screening programs. Should any of these remaining districts initiate early screening programs, it is likely that variations in instruments would be increased over present levels.

Most school districts reporting early screening programs practiced universal screening, i.e., 100 percent of the children were screened at one or more grade levels. Seven of 177 school districts practicing screening did so selectively, targeting the program to fewer than all children. However, further variability was noted in the levels at which screening occured. Screening at "kindergarten only" was practiced in 83 school districts; kindergarten and prekindergarten in 31 school districts; "prekindergarten only" in 28 school districts; kindergarten and grade one in 18 school districts; "grade one only" in 11 school districts; and all three levels in 6 school districts.

Twelve learning impediments were specified in the New York State survey and school districts having screening programs were asked to identify the tests they used in each of these categories. Again ex-

treme variability was noted. The following details the 12 learning impediments and the number of different instruments reportedly used in screening for them:

Condition	# of different instruments used
learning disabilities	152
mental retardation	59
brain damage	57
emotional disturbance	38
cultural disadvantage	47
general behavior	104
sensory motor integration	132
laterality and directionality	104
visual and auditory perception/acuity	129
conceptual skill	113
language development	111
previous academic experience	50

Consequences of Extreme Variability in Screening Programs

School districts that establish early childhood screening programs will, in the long run, identify greater numbers of children with learning impediments proportional to their total enrollment. Many special education funding policies link the level of financial support to the number of children identified and served by a school district. For example, Public Law 94-142 of the 94th Congress of the United State of America bases maximum authorized funding on the number of handicapped children aged three to twenty-one, inclusive, who are receiving special education and related services. An additional sum of 300 dollars per handicapped child aged three to five who has been identified and is being served is provided for. Since fifty to seventy-five percent of

these funds, when allocated, are to pass through to LEAs, schools that have identified greater numbers of children with learning impediments will obtain increased levels of support.

To the extent that rural and suburban schools improve their identification and service components for children with learning impediments, and urban school districts fail to do so, maldistributions of resources over the state as a whole are likely to become exacerbated. Any funding policies linked to numbers of students with learning impediments served will transfer more funds to school districts with good screening programs.

The proportion of children identified as "high risk" in screening will vary according to the breadth of the screening program. Universal screening policies will detect greater numbers of children in "high risk" categories and so will screenings that occur on more than one occasion. Also, screening batteries that include a wider array of tests and sample a larger number of performance areas will identify more children with learning impediments. Therefore, incidence rates for various learning impediments within school districts will be influenced by the sophistication of the initial screening procedures. In contrast, narrow screening processes will result in greater numbers of "false negatives."

Any variations in the instruments used for identifying children with learning impediments will be reflected in the statewide school population statistics and therefore will affect the planning base. Since most of the tests presently used in New York State yield data that is logically incomparable, inferential distortions are likely to

emerge if these data are treated simultaneously in an analysis. Wide instrument variations provide a very inadequate base for making comparisons between school districts or constructing a statistical profile of the state as a whole.

Extreme variability in screening procedures among school districts also results in redundancies and inefficiencies growing out of duplication of effort in organizing and administering screening programs. In selecting screening instruments, each school district may expend many valuable professional manhours. If they also design their own screening devices, the professional cost is greater and the duplication of effort more pronounced.

A further negative consequence of extreme variability is the hindering of the general growth of competence among the teachers of the state in the area of screening, administration and interpretation. Teachers who are mobile between school districts will require a period of learning to become adequate in the screening techniques of the new district. Extreme variability in screening procedures also precludes the development of inservice teacher education packages that are suitable for statewide dissemination.

Extreme variations in instruments used within a learning impediment category, such as learning disabilities, allow school districts to arrive at idiosyncratic definitions of learning impediments which may be inconsistent with the best available present knowledge. Furthermore, these idiosyncracies may result in a child being suspected of "learning disability" in one school district and "mental retardation" on another. Failures on similar screening tasks may indicate different conditions

depending on the category which an instrument is assigned to detect by the local school district. For example, some school districts in New York State view failure on the Developmental Test of Visual Perception indicative of mental retardation, others view it as indicative of cultural disadvantagement, or general behavior problems, and some even, language development.

In general, extreme variability leads to poor accountability because each school district provides its own separate standard of excellence, standards which remain incomparable between districts. Lacking consistent bases, both in the form of policy and procedures relating to screening, it becomes difficult to comprehensively evaluate the adequacy of the system as a whole or any of its elements. Without the ability to do this, it becomes extremely difficult to target areas for improvement, to establish general goals, or to document change.

PART V

PUBLISHED INSTRUMENTS USED IN NEW YORK STATE SCREENING

The following section provides the reader with a technical overview of the published tests reportedly used in New York early screeening programs. To enable meaningful analysis, the 30 most frequently used published instruments were reviewed in detail. It is worth noting that the range of instances of application for particular instruments ranged from 399 (Metropolitan Readiness) to six (Riley Preschool Development Screening Inventory). In other words, the thirty instruments selected for analysis virtually exhausted the array of published instruments used for screening anywhere in New York State.

First, the reader's attention is directed to Table 1. Contained in that table are frequencies of application for tests by impediment category or condition. Appropriate applications are tabulated using numbers in bolder typeface. Also, an asterisk (*) has been assigned to further draw attention to appropriate instrument applications. An "appropriate application" only means that, on the basis of the item content, face validity, or administrative manual, the instrument might detect the presence of a particular condition. Many of the instruments that meet this liberal or minimal standard remain totally inadequate in validity and technical excellence. An analysis of instrument adequacy appears in Tables 2 and 3.

In brief, what does Table 1 reveal?

1. Of 2,059 reported application of screening instruments in New York State, 1,069 displayed some degree of compatibility between test content and learning impediment screened for. This represents a 52% appropriate application rate.

2. Rates of appropriate screening instrument application for each of the learning impediments were as follows:

a.	learning disabilities	71%
b.	mental retardation	59%
c.	brain damage	21%
d.	emotional disturbance	30%
e.	cultural disadvantage	16%
f.	general behavior	25%
g.	sensory-motor integration	82%
h.	laterality/directionality	15%
i.	visual/auditory	75%
j.	conceptual skills	20%
l.	language development	71%
m.	previous academic experience	44%

Table 2 merely provides a summary of the conclusions that were reached by the Center for the Study of Evaluation under the direction of Ralph Hoepfner and Stephen Klein of UCLA regarding the adequacies of preschool and kindergarten assessment devices. Not all tests reportedly used in New York State were reviewed by UCLA because some of them are of recent vintage and others served very specialized functions and therefore are not widely disseminated within the educational community.

The data in Table 2 tells us something important, however. It is very difficult to identify a test which meets basic standards of excellence. This statement especially holds true for instruments used with young children. When tests lack technical excellence, a good deal of the individual differences that we detect through their use have little to do with variations in the underlying entities we wish to measure; rather they are likely to be function of variation on unwanted factors or random error influences. Many professional man-hours are dissipated interpreting random variation among test results which are a consequence of the tests' inadequacies.

Table 3 evaluates 30 of the most widely used tests in New York State with respect to suitability for detecting the 12 learning impediments.

Table 1. Published Assessment Devices Most Frequently Used by New York State Schools in Early Educational Screening Programs for Learning Impediments: 177 of 736 School districts Reporting Screening Programs (1973-74)

ASSESSMENT DEVICE	Frequency of Use for Detecting Condition:*												Total Applications
	Learning Disabil.	Men.ret.	Brain d.	Em. dist.	Cul.dist.	Gen.beh.	SM-integ.	Lat./dir.	Vis./aud. Percep/Ac	Con.skl	Lan.dev.	Pr. ac.	
1. Metropolitan Readiness	104*	20*	25	17	31	21	29*	19	30*	37	33*	26*	399
2. Vane Kindergarten Test	83*	23*	27	20	12	22*	13*	20	25*	19	25	13	302
3. Boehm Test of Basic Concepts	38*	1*	4		12*	2	8	5*	2	15*	14*	6	107
4. Devel. Test of Visual Percep.	36	8	3		8	5	14*	9	15	6	5		109
5. ABC Inventory	29					4	4*	4	4*	5	5	2	57
6. Gesell Developmental Scheds	27*		5			8*	8*	5	7*	6	3*	1	70
7. Ill. Test of Psycholinguistic Abilities	18*	2	2			1	2*	2	7*	4*	7*		45
8. Pupil Rating Scale for LD	18*			7		4	4*	5*	1*	4	4*	2	49
9. Bender-Gestalt Test	16*	6	4*	4*	4	5	9*		7*	1	7		63
10. Wepman Auditory Discrim. Test	15*	6			6	5	6	5	5*	6	4		59
11. Purdue Perceptual Motor Surv.	15*		4				10*	10*	1*				40
12. Human Figure Drawings	39	22*	12*	22*	20	20	28*	18*	22	22	7	4	236
13. Rutgers Drawing Test	6		2				5		3				16
14. Coop. Presch. Inventory Test	12	7		7		4	4*	4	4	4*	3		49
15. WISC	12	5		5		1	5		1	4	4	4	41
16. Stanford Binet	5	6		5	4	3	3			3	3	3	35
17. Denver Developm. Screening T.	13*		4			1	5*	1	4		5*		33
18. First Grade Screening T.	11	4*	4*		4	1	6*	5	1*	4*	1*		41
19. Wide Range Achievement T.	6	6		6		4	5		1	4	4*	4*	40
20. Screening Tests for SLD (Slingerland)	16*				4	2	4*	9	7*	5	4*	1	52
21. Meeting Street	22*					1	5*	4	5*	2		1	40
22. Perceptual Forms Test (Winterhaven)	27*	2	5			2	9*	5	9*	2	3		64
23. Otis-Lennon Mental Ability T.					7*	2	2	2	2*	2*	2*		21
24. Dev. Test of Vis. Mot. Integra	9		3			2	3*	3		1			21
25. White Phillips	10	4											14
26. Stanford Early Achievement T.	4	4*							3*	3*		3*	17
27. SRA Primary Mental Abilities			3		4				2*	2	4*	4	19
28. Ilg & Ames Kindergarten Screening	6	1*				1	1*		1*	1	3		14

**See Page Two for continuation of Table 1.

Table 1. Published Assessment Devices Most Frequently Used by New York State Schools in Early Educational Screening Programs for Learning Impediments: 177 of 736 School Districts Reporting Screening Programs (1973-74)

ASSESSMENT DEVICE	Frequency of Use for Detecting Condition:*												Total Applications
	Learning Disabil	Men. ret.	Brain d.	Em. Dist.	Cul. dist.	Gen. beh.	Sm-integ.	Lat/dir.	Vis/aud. Percep/Ac	Con. skl.	Lan. dev.	Pr. ac.	
29. PPVT	1	2*									3*		6
30. Riley Presch. Dev. Screen, Invt.			3*	3*									
Total Number of Different Devices Used to Assess Condition	152	59	57	38	47	104	132	104	129	113	111	50	

*** 1. Learning disability; 2. mental retardation; 3. brain damage; 4. emotional disturbance; 5. cultural disadvantage; 6. general behavior; 7. sensory-motor integration; 8. lateraity/directionality; 9. visual-auditory perception and acuity; 10. conceptual skills; 11. language development; 12. previous academic experience.

Table 2. Summaries of Adequacy Evaluations of Major Published Tests Used in New York State Screening: Center for the Study of Evaluation and Early Childhood Research Center, UCLA

TEST NAME Educational Objective		Measurement Validity	Examinee Appropriateness	Admin. Validity	Normed Technical Excellence
GESELL DEVELOPMENTAL SCHEDULES					
Personal-Social		Poor	_____	Poor	Poor
Language		Fair	_____	Poor	Poor
Motor		Poor	_____	Poor	Poor
Adaptive Behavior		Poor	_____	Poor	Poor
BOEHM TEST OF BASIC CONCEPTS		Fair	Good	Good	Fair
MARIANNE FROSTIG DEVELOPMENTAL TEST OF VISUAL PERCEPTION					
Constancy of Shape	Pre-K	Poor	Fair	Poor	Poor
	Kdg.	Poor	Fair	Fair	Poor
Eye-Motor Coord.	Pre-K	Fair	Fair	Poor	Poor
	Kdg.	Fair	Good	Fair	Poor
Position in Space	Pre-K	Poor	Fair	Fair	Poor
	Kdg.	Poor	Good	Fair	Fair
Spatial Relations	Pre-K	Fair	Fair	Poor	Poor
	Kdg.	Fair	Fair	Fair	Poor
Total	Pre-K	Poor	Fair	Poor	Poor
	Kdg.	Poor	Fair	Poor	Poor
ABC INVENTORY					
Total		Poor	Fair	Fair	Poor
METROPOLITAN READINESS (KDG)					
Alphabet		Fair	Good	Good	Fair
Copying		Good	Good	Fair	Poor
Listening		Fair	Good	Good	Poor
Matching		Fair	Good	Good	Poor
Numbers		Good	Fair	Good	Poor
Word Meaning		Good	Good	Good	Poor
Total		Fair	Fair	Fair	Fair
BENDER VISUAL MOTOR GESTALT					
Total	Pre-K	Fair	Fair	Poor	Poor
	Kdg.	Fair	Fair	Poor	Poor
WEPMAN AUDITORY DISCRIMINATION (KDG)					
Total		Fair	Good	Fair	Poor

Table 2 cont'd

TEST NAME Educational Objective		Measurement Validity	Examinee Appropriateness	Admin. Validity	Normed Technical Excellence
GOODENOUGH Harris Drawing Test					
Point Scale: Man	Pre-K	Poor	Good	Poor	Poor
	Kdg.	Poor	Good	Fair	Poor
Point Scale: Self	Pre-K	Poor	Good	Poor	Poor
	Kdg.	Poor	Good	Fair	Poor
Point Scale: Woman	Pre-K	Poor	Good	Fair	Poor
	Kdg.	Poor	Good	Fair	Poor
Quality Scale: Man	Pre-K	Poor	Good	Poor	Poor
	Kdg.	Poor	Good	Fair	Poor
RUTGERS DRAWING TEST					
Total	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Good	Fair	Poor
CALDWELL PRESCHOOL INVENTORY					
Assoc. Vocab.	Pre-K	Fair	Fair	Poor	Poor
	Kdg.	Fair	Fair	Poor	Poor
Concept Activation	Pre-K	Poor	Fair	Poor	Poor
Numberical	Kdg.	Poor	Fair	Poor	Poor
Concept Activation	Pre-K	Poor	Fair	Poor	Poor
Sensory	Kdg.	Poor	Fair	Poor	Poor
Personal-Social	Pre-K	Poor	Fair	Poor	Poor
Responsiveness	Kdg.	Poor	Fair	Poor	Poor
Total	Pre-K	Poor	Fair	Fair	Fair
	Kdg.	Poor	Fair	Fair	Fair
WIDE RANGE ACHIEVEMENT TEST (Kdg.)					
Arithemetic		Fair	Fair	Fair	Fair
Reading		Fair	Fair	Fair	Fair
Spelling		Fair	Poor	Fair	Fair
WINTER HAVEN PERCEPTUAL TESTING (First Grade Test)					
Total Test		Poor	Good	Poor	Poor
OTIS-LENNON MENTAL ABILITIES TEST					
Total Test		Poor	Fair	Good	Fair
DEVELOPMENTAL TEST OF VISUAL - MOTOR INTEGRATION (Beery-Buktenica)					
Total Test	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor
PEABODY PICTURE VOCABULARY TEST					
Total Test	Pre-K	Good	Fair	Fair	Fair
	Kdg.	Good	Fair	Fair	Fair

Table 2 cont'd

TEST NAME Sub-Test		Measurement Validity	Examinee Appropriateness	Admin. Validity	Normed Technical Excellence
STANFORD EARLY SCHOOL ACHIEVEMENT TEST (Kdg.)					
Aural Comprehension		Poor	Fair	Good	Poor
Environment		Fair	Fair	Good	Fair
Letters and Sounds		Fair	Fair	Good	Poor
Mathematics		Poor	Fair	Good	Poor
Total		Poor	Fair	Fair	Poor
PRIMARY MENTAL ABILITIES (Kdg.)					
Number Facility		Poor	Fair	Fair	Poor
Perceptual Speed		Fair	Good	Fair	Poor
Spatial Relations		Fair	Fair	Fair	Poor
Verbal Meaning		Fair	Fair	Fair	Poor
Total		Fair	Fair	Fair	Poor
ILLINOIS TEST OF PSYCHOLINGUSTIC ABILITY					
Auditory Association	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Poor	Good	Fair	Poor
Auditory Closure	Pre-K	Poor	Fair	Fair	Poor
	Kdg.	Poor	Fair	Fair	Poor
Auditory Reception	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Good	Fair	Poor
Aud. Seq. Memory	Pre-K	Fair	Good	Fair	Poor
	Kdg.	Fair	Good	Fair	Poor
Manual Regression	Pre-K	Poor	Good	Fair	Poor
	Kdg.	Poor	Good	Fair	Poor
Sound Blending	Pre-K	Poor	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor
Verbal Regression	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor
Visual Association	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor
Visual Closure	Pre-K	Poor	Fair	Fair	Poor
	Kdg.	Poor	Fair	Fair	Poor
Visual Reception	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor
Visual Seq. Mem.	Pre-K	Fair	Fair	Fair	Poor
	Kdg.	Fair	Fair	Fair	Poor

Table 3. Evaluations of Published Tests Used in New York State Screening: Multiple Standards

TEST NAME	POSSIBLY SUITABLE FOR DETECTING												Teacher administer	Pre-K	Kdg.	Grade 1 +	Group test	Spanish Version	Buros reliability	Buros validity	Rated "fair" or better in 3/4 UCLA criteria	Screening instrument
	LD	MR	BD	ED	CD	GB	SMI	Lat. Dir.	V/A Per.	Cog. Skl.	Lan. Dev.	Prv. Acd.										
A. Tests Unsuited for Screening																						
1. Gesell Developmental Schedls	+	+				+	+		+		+		no	+	+				poor	poor	no	no
2. Illinois Test of Psycho-linguistic Abilities	+						+		+	+	+		?	+	+	+			fair	poor	no	no
3. Rutgers Drawing Test	?	?		+	?		+		+V					+	+	+	?	?	fair	fair	unrated	no
4. Stanford Binet	+	+	widely used individual intelligence test unsuited for screening																			no
5. Vane Kindergarten Test	+	+	?	?		+	+		+V				no		+	+			fair	fair	unrated	no
6. WISC	+	+	widely used individual intelligence test unsuited for screening																			
B. Tests Suited For Screening																						
7. ABC Inventory							+			?	?		+	+	+	+			poor	poor	no	yes
8. Bender-Gestalt Test	+		+	+					+V					+	+	+	?		poor	fair	no	no
9. Boehm Test of Basic Concepts	+	+			+			+		+	+		+		+	+	+	+	good	fair	yes	yes
10. Coop. Pre-sch. Inventory Test			+		+		+			+		+	+	+	+	+			poor	poor	no	yes
11. Denver Developmental Screening Test	+	+			+		+			+	+		+		+	+	+		good	fair	unrated	yes
12. Dev. Test of Vis.Mot.Integra.							+		+V				+		+	+	+		fair	poor	yes	yes
13. Developmental Test of Visual Perception	?		?				+		+V				+	+	+	+	+		fair	fair	unrated	yes
14. First Grade Screening Test		+	+	+					+AV	+	+	+	+		+	+	+		?	fair	no	yes

Table 3. Evaluations of Published Tests Used in New York State Screening: Multiple Standards

TEST NAME	POSSIBLY SUITABLE FOR DETECTING													Teacher administer	Pre-K	Kdg.	Grade 1 +	Group test	Spanish Version	Buros reliability	Buros validity	Rated "fair" or better in 3/4 UCLA criteria	Screening Instrument
	LD	MR	BD	ED	CD	CB	SMI	Lat. Dir.	V/A Per.	Cog. Skl.	Lan. Dev.	Prv. Acd.											
15. Human Figure Drawings		+	+	+			+							?	+	+	+	+		fair	fair	no	yes
16. Ilg & Ames Kindg.Screening		+					+		+AV					?		+	+			?	?	no	yes
17. Meeting Street	+						+		+ V		+			+		+	+			poor	poor	unrated	yes
18. Metropolitan Readiness	+	+					+		+ V		+	+		+		+		+		good	good	yes	yes
19. Otis-Lennon Mental Abil.Tst		+			?				+ V	+	+			+		+	+	+		good	good	yes	yes
20. Peabody Picture Vocab.Test		+			?						+			+	+	+	+		+	fair	fair	yes	yes
21. Perceptual Forms Test (Winterhaven)	?						+		+ V					+		+	+	+		poor	poor	unrated	yes
22. Pupil Rating Scale for LD	+		+			+	+	+	+ A		+			+	+	+	+			unrated		unrated	yes
23. Purdue Perceptl Motor Survey																							
24. Riley Presch.Dev.Screeng. Inventory		+	+	+			+							+	+	+				?	?	unrated	yes
25. Screening Test for SLD (Slingerland)	+						+		+AV		+			?				+	+	poor	poor	unrated	yes
26. SRA Primary Mental Abilities	?	+					+		+ V		+				+	+	+			fair	poor	yes	yes
27. Stanford Early Achievement Tst.		+			?				+A	+	+	+		+	+	+	+			fair	poor	no	yes
28. Wepman Auditory Discrim. Test	+								+A					+	+	+				fair	fair	yes	yes
29. White Phillips																							
30. Wide Range Achievement Test											+	+		+	+	+				fair	poor	yes	?

PART VI

THE COMBINED PUBLIC AWARENESS AND FORMAL SCREENING OPTION

One of the basic strategies in the early identification of children with learning impediments is to initiate a broad and effective public awareness program. Clearly a public awareness program must be organized in such a way as to insure that messages are communicated in a manner that is understandable to all language, ethnic, and cultural groups.

The Southwest Illinois Regional Special Education Association, the Belleville Area Special Education Districts, and Dr. William Whiteside of Edwardsville have developed strategies for meeting these needs that may be effectively exported. Included among them are the following:

1. The school system may host a meeting of concerned leaders, representing a cross section of the community, who will be able to advise the school system of the various groups within the community needing special consideration in terms of language, ethnic and/or cultural considerations.

2. The school district may consider establishing a permanent local advisory committee for the same purposes as listed above, which will advise and assist in preparation of such materials. Individuals on the committee may assist and/or recommend others who can help with the writing of materials in languages other than English.

3. The school system, or either of the above groups, should request that a local survey be taken to document the number of persons of families represented in each language, ethnic, and/or cultural group. Such a list would benefit the determination of the various ethnic and cultural groups with which you wish to communicate. When preparing the list, provisions should be made for English speaking handicapped children whose parents speak a different language.

In addition to enlisting the support of families whose children may have special educational or developmental needs, a public awareness program should inform parents about the availability of services, the human rights

and legal rights of exceptional children, and what steps to take in order to initiate action on a problem. The following represent specific public awareness activities which serve, in many instances, as a latent screening function.

1. If the district has a newsletter which is mailed to each parent of a school child, a special issue may be printed to describe the special education program and the rights of handicapped pupils. If the district has no such newsletter, creation of one for this purpose should be considered.

2. The traditional "send a note home with the kids" has not been found to be particularly effective. If this distribution method is utilized, a signed return slip should be provided. The district could then mail notices to those parents whose children did not return the signed slip.

3. Notices, in the form of special interest ads or in feature articles, can be placed in both metropolitan and local newspapers.

4. Spot announcements on radio and TV or short features on newsworthy educational programs may be used.

5. During Exceptional Children's Week in the month of May, extra effort should be exerted.

6. Special displays created by regional materials centers, by local media people, or classes may be displayed in local shopping centers.

7. Teachers and other school personnel may volunteer as speakers for various service and social organizations. The local district should make their services known to these groups. Slide presentations are very effective speaker aids.

8. One or more civic or social groups may take on the publication of services for the handicapped as their special interest project for the year.

9. The district may establish a Special Education Day in the community, including such activities as an open house, guided tours, and teas.

10. The School District may create a "Special Education Program of the Month" to focus attention on specific services.

11. The local district may create an award of the month for various persons within the community who have assisted exceptional students and their programs within the public schools.

12. Someone within the community may establish a special education flea market with items for sale that may be used by numerous types of handicapped children, such as: clothing, puzzles, games, walkers, wheelchairs, magnifiers, and special furniture. The local district should also have an information booth with printed material available.

13. The County Advisory Committee may hold hearings, and information sessions in communities throughout the country.

14. An Arts and Crafts Fair for the handicapped could be held featuring exhibits which include the handicraft of exceptional children, professional personnel working with these children, and others interested in exceptional children. Space rental or a percentage of sales could be donated for equipment and materials for special classes. Again the district could utilize the media to publicize the event and disseminate pertinent information.

15. A number of these activities can be used separately or in combination at the various local and county homecomings in the community.

Location of Exceptional Children

Many public awareness activities will also assist in the location of exceptional children. The following suggestions, therefore, will specifically aid the process of locating these children.

1. The school district may convene a meeting of representatives from those state, county, local and/or parents groups concerned with the education and welfare of exceptional individuals. The purpose of such a meeting would be to gather suggestions and assistance with the location of handicapped inidviduals between the ages of three and twenty-one who reside in the community and may need special educational services. Such a meeting would also avoid duplication of services while preventing service gaps.

2. Annually, screening may be conducted for exceptional children. Such screening may be conducted via the suggestions made under public awareness. In addition, special contacts and/or information may be provided by various medical personnel, private pre-schools, day care centers, or social agencies most likely to have contact with pre-school children.

3. Hearing and vision screening should be conducted at regular intervals. The time of such screenings should be provided to parents within the district and an invitation extended to have them bring in other children in the family that they would like to have screened.

4. Speech and language screening is to be done upon initial enrollment in the district and at such times as it appears necessary.

5. Annually, all teachers in the district, and other professional personnel, should screen the children with whom they have contact, and make referrals to the appropriate sources.

6. All types of screening may be conducted, on an ongoing and special events basis, to encourage the parents and members of the community having children to look to the local schools as a source of assistance if they suspect that a child may have any type of problems which may interfere with his educational progress either now or in the future.

A Recommended Battery of Screening Instruments

If the state of New York adheres to the taxonomy of learning impediments around which data for the present report was gathered, and if our selection

process is confined to the thirty published instruments previously discussed, the following battery of tests would provide some preliminary information concerning these impediments:

Pre-Kindergarten Level:

1. Pupil Rating Scale for Learning Disabilities
2. Draw-A-Person (using the Kippitz scoring system)

Kindergarten Level:

1. Boehm Test of Basic Concepts (Spanish version available)
2. Wepman Auditory Discrimination Test
3. Metropolitan Readiness Test

The above battery allows for information to be generated at both pre-kindergarten and kindergarten levels. All tests are suitable for teacher administration and scoring while also meeting above average standards of technical excellence. Only one of the five requires individual administration: The Wepman Auditory Discrimination Test. Teacher observations and ratings are also included in the screening battery via the Pupil Rating Scale for Learning Disabilities. Provision would be made for some testing to be conducted in Spanish with an instrument of known psychometric characteristics. In addition, the battery would capitalize on common usage and familiarity by educators.

In the opinion of the author, the above battery would meet cost-efficiency standards, but would require in-service education for teachers and careful attention to the sequencing of testing during pre-kindergarten and kindergarten.

PART VII

OVERVIEWS OF NEW YORK STATE SCREENING INSTRUMENTS

The test overviews contained in this section of the report were derived from:

1. Examination of the instruments and their manuals;
2. Reviews of criticisms of the instruments published in Buros, The Seventh Mental Measurement Yearbook;
3. Interviews with those having direct experience using many of the tests in educational settings.

Discussions of the WISC and Standford Binet were excluded because of their irrelevance for screening purposes. These are major intelligence tests used by psychologists as part of a diagnostic or case study process.

The ABC Inventory[1]

The ABC Inventory purportedly measures kindergarten and school readiness, identifying children who are developmentally immature. Readiness scores from 3-6 to 6-7 years are derived. Tasks which comprise the instrument are:

1. Drawing a man
2. Characteristics of objects
3. General information
4. Simple numbers and shapes

While instructions for administration are clear and brief, and the test possesses some content validity, the following areas are profoundly lacking in excellence:

1. Information on standardization group;
2. Information on reliability;
3. Information on predictive validity (see the Metropolitan Readiness Test for examples of good predictive validity).

There is little evidence to support the use of this test as part of a screening program and in fact some of the information printed in the test manual (only 3 pages) is misleading and contrary to American Psychological Association standards.

[1]The ABC Inventory, Normand Adair and George Blesch, Research Concepts.

Bender-Gestalt Test[2]

The Bender-Gestalt Test is a clinical/diagnostic tool useful with children ages four and over. According to the School Evaluation Project, Center for the Study of Evaluation, UCLA, the Bender measures the educational objective area of eye-hand coordination. However, in clinical applications the test has been used to detect organicity, and to reveal something about personality dynamics and developmental level.

Reporting that the Bender is used in an assessment program in the schools is rather imprecise because there are several versions of the test:

1. Visual Motor Gestalt Test (VMGT) - Stimulus pictures are available in 35 mm slides to allow group administration.
2. The Hutt Adaptation of the Bender-Gestalt Test (HABGT) - Modifications in designs and drawing method.
3. The Bender Visual Motor Gestalt Test For Children - Ages 7-11
4. The Bender Gestalt Test For Young Children 5-10 - Uses a developmental scoring system by Koppitz.

Different versions of the test have been used with different age groups and various scoring systems address themselves to developmental level, organicity, and emotional indicators. As such, the test has been used both projectively and non-projectively.

The Bender-Gestalt Test has demonstrated sufficient unreliability to cause mention in the instruction booklet of factors that may cause fluctuations in performance. Many details of the administration have been shown to influence test outcomes, e.g., allowing examinee to move materials from their original position, number of designs to be copied, presence or absence of verbal prompting to perform as well as possible.

While unsuited for use in screening programs, the Bender "should be included, if possible, in every diagnostic examination of children from age five because of its unique contributions to the evolution of perceptual-motor functioning, neurological impairment, expressive styles, and maladjustment." (Kitay, Philip. "The Bender-Gestalt Test" in Buros, The Seventh Mental Measurement Yearbook, 394-395).

[2]The Bender-Gestalt Test, Gerald Pascal and Barbara Suttell, Grune and Stratton, Inc.

Boehm Test of Basic Concepts[3]

The Boehm Test of Basic Concepts is a small group test which is suited for teacher administration and scoring with children of grades kindergarten through two. Fifty items are included in the test, broken down into two booklets for administration in two sessions. Responses to items involve making x's on pictures. The pictures are large and include illustrations of minority children.

According to the author, the purpose of the test is "to assess beginning school children's knowledge of frequently used basic concepts widely, but sometimes mistakenly, assumed to be familiar to children at their time of entry into kindergarten or first grade." Generally assessed by the Boehm are:

1. Understanding of and following directions (spoken)
2. Understanding of space (direction, orientation, dimensions)
3. Understanding of time and quantity

A strength of the test is that the results are readily translated into remedial plans, and the test can be used to trace changes in performance resulting from instruction.

The validity of the test has generally been examined only in the area of item content, and criticisms have been raised because of the absence of other kinds of validity studies. Reliability information and discussion of test error are well presented in the accompanying test manual. Item selection was conducted according to psychometric principles.

The normative sample includes urban, rural, midwestern, southeastern, and northeastern schools. A large percentage of the lower SES children in the normative sample were enrolled in schools in the southeast. No assertion is made that the normative sample is representative of the U.S. as a whole.

In comparison with most other kindergarten-level tests, the Boehm rates high.

[3]Boehm Test of Basic Concepts (1968), Ann E. Boehm, Psychological Corporation.

Cooperative Preschool Inventory (Caldwell)[4]

The Caldwell Cooperative Preschool Inventory (CPI) was designed for assessment and screening of disadvantaged children ages three to six. Individually administered by a classroom teacher in about 15 minutes (64 items) the CPI involves:

1. Labelling quantities
2. Identifying serial positions
3. Perceiving shapes
4. Executing basic drawing skills

Factors which the test is designed to measure are:

1. Associative vocabulary
2. Concept activation (numerical)
3. Concept activation (sensory)
4. Personal-social regressiveness

Educational objective areas assessed by the test include, according to the Center for the Study of Evaluation, UCLA:

1. Functioning and Structure of the Human Body
2. Counting and Operations with Integers
3. Oral Semantic Skills
4. Visual Discrimination and Recognition

Standardization of the test occurred in eleven Head Start centers with only English speaking subjects. Technical data is available for the 3-0 to 6-5 age groups. From this limited normative sample, reliability coefficients (internal consistency) from .86 to .92 were reported. Data on the predictive validity of the test are unavailable.

Instructions for administration and scoring are clear, but it has been suggested that the test was prematurely marketed with insufficient normative and technical data.

[4]Cooperative Preschool Inventory, Revised, Bettye Caldwell, Cooperative Tests and Services.

Denver Developmental Screening Test[5]

The Denver Developmental Screening Test (DDST), developed by William K. Frankenburg and Josiah B. Dodds, aims to be a simple, clinically useful tool for the detection of serious developmental delays. The test was developed for use by persons who lack formal training in psychological testing. It evaluates four aspects of a child's functioning: gross motor, fine motor, adaptive language, and personal-social development.

The test consists of 105 items which were administered to a standardization population of 1,036 children between the ages of two weeks and 6.4 years. Though a relatively small number of children were retested by four separate examiners for reliability purposes, agreement between 90 and 100 per cent on each of the items, arbitrarily assigned to the four measured areas of functioning, appeared to justify the authors' claims for a high level of reliability in administration and scoring. A parallel study produced a correlation of .97, which suggests high validity, especially since there were few cases of discrepancy between the two tests in identifying normal and deviant children. The rates of 3.2 per cent overreferrals and 0.4 per cent underreferrals are considered to be highly satisfactory for any screening instrument.

There are some weaknesses in DDST with regard to standardization, reliability, and validity. The Menninger Foundation reported in a review: "From clinical use of the DDST we have some reason to question the circumscribed geographical selection of the sample and applicability of the norms in screening children from the lower socioeconomic groups, especially in the language area." The norm group contained a significantly higher proportion of white children and of children whose fathers

are in the professional, managerial, or sales occupations than the census distribution would warrant. Test-retest reliabilities are based on a extremely small sample of children, twenty to be exact, although a more recent test-retest reliability study of 186 children produced coefficients between .66 and .93, with the coefficient increasing as the age increased; the best results are after 30 months of age.

Despite these weaknesses, the DDST has proven itself to be a practical efficient, and dependable device. It is inexpensive, quick, and easy to administer and evaluate with relatively little training or experience in testing. The manual is direct and clear, and scoring guides are explicit. The materials themselves are colorful and of inherent interest to young children. Although the available evidence indicates that it is not as reliable, valid, and sensitive as its authors had hoped, it does meet acceptable standards of reliability and validity. According to Frankenburg, Camp, and Van Natta, the test consistently identifies children with low IQ's and does not identify as "deviant" those children who are actually normal. This is especially encouraging at the upper ages. With the reservations stated in mind, along with the fact that a screening tool is only as good as the sensitivity of its user to the behavior of young children, the DDST, in the context of its stated aims to provide screening of developmental delay in young children, is an efficient and predictive screening instrument.

[5]Denver Developmental Screening Test, William Frankenburg and Josiah Dodds, Ladaca Project and Publishing Foundation, Inc.

Developmental Test of Visual-Motor Integration (Beery-Buktenica)[6]

The Developmental Test of Visual-Motor Integration (DTVMI), often referred to as "the Beery," is a geometric forms copying test for children ages two to 15. Teachers can administer the test in a group situation. A short form (15 items) and a long form (24 items) is available.

The DTVMI is among the better geometric form copying tests. Special attention is given to the design of the test booklet to circumvent problems commonly encountered in tests of this type i.e. spareness of paper and positioning of the stimulus items. Administration procedures are clear but scoring entails considerable subjectivity. Item selection and sequencing, however, are based on firm theoretical considerations and are rated good.

A number of reliability studies have been made, with small samples, and both internal consistency and stability results appear satisfactory. Evidence on the predictive validity of the test is lacking. Correlations with chronological age, mental age and reading achievement have been presented as indications of test validity.

The test was standardized on 1,039 Illinois children with a considerable over-representation of suburban children. Average children were selected to yield an average performance standard at different ages.

[6]Developmental Test of Visual-Motor Integration, Keith Beery and Norman Buktenica, Follett Educational Corporation

First Grade Screening Test[7]

The First Grade Screening Test is a group test suited for teacher administration at the end of kindergarten or the beginning of first grade. Boys and girls receive different kinds of test booklets to "allow children an opportunity to identify with familiar though different social rules." Although scoring procedures do not result in any differential diagnosis, the test is sensitive to intellectual deficiency, central nervous system dysfunctions and emotional disturbances. No time limits are specified for the test; however, it typically takes up to 45 minutes to complete.

In the test manual it is asserted that the purpose of the test is to: screen potential or beginning first-grade students in order to identify the few who will not, without benefit of special assistance, make sufficient progress to be ready for second grade the next year. While two levels of norms are provided in the manual (end of kindergarten; beginning of first grade), schools are encouraged to develop local norms which take into account:

1. Goals of first-grade instruction
2. Teacher-pupil ratio
3. Availability of diagnostic resources
4. Availability of services for handicapped

Psychometric analysis of the twenty-two items in the test reveals fair reliability. While content validity of the test appears high, its usefulness as a predictive device has yet to be established.

Items included in the test are:

1. Human figure and geometric drawings

2. Eye-hand coordination

3. Identifying objects

4. Personal/social content items

5. Self-concept

6. Vocabulary - receptive

7. Cognition of figural units

8. Memory - short term

[7] First Grade Screening Test, John E. Pate and Warren W. Webb, American Guidance Services Inc.

Gesell Developmental Schedules[8**]

Gesell Developmental Schedules (GDS) are observational schedules used with children of ages four weeks to 72 months. These schedules include a number of criteria measures organized according to developmental level in three volumes: The Psychology of Early Growth; The First Five Years of Life; and Developmental Diagnosis. Frequently used by psychologists, pediatricians, and neurologists, the GSD are subjective in scoring and their realiability will vary in accordance with the skills of the examiner. These schedules are used individually with children, require a considerable amount of time to administer and score, and cannot be effectively used by the average teacher.

Validities of the schedules are weak or indeterminant. Early records show little association with subsequent performance. Some of the correlations between results from the Gesell Developmental Schedules and the Stanford Binet can be accounted for by the fact that the GDS include some items taken from the Stanford Binet.

The schedules allow for analysis of:

1. Postural behavior
2. Prehensory behavior
3. Perceptual behavior
4. Adaptive behavior
5. Language - social behavior

Educational objectives for which the Center for the Study of Education lists the Gesell Developmental Schedules as being targeted toward include:

1. Development of Personality (dependence-independence)

2. Physical Coordination

3. Eye-Hand Coordination

4. Oral Semantic Skills

The schedules are considered in need of revision both in content and scoring standards.

[8]Gesell Developmental Schedules (1925-49), Arnold Gesell and others, Psychological Corportation.

**Also, see Ilg and Ames Kindergarten Screening Test.

Human Figure Drawing Tests[9]

Among the screening devices reportedly used in New York State's school districts, various human figure drawing tests rank first in frequency of application (see Table 1). Responses to the 1973 statewide survey revealed that human figure drawing tests were used as a means of assessing every type of learning impediment. No learning impediment was identified in the survey where some school district in New York did not use a HFD test for screening. The lowest reported rate of application in screening was for assessing previous academic experience (four school districts) while the highest rate was for detecting learning disabilities (39 school districts).

Many of the New York State applications are questionable because the use of human figure drawings (HFD) is mainly limited to two major uses in assessing children: first, to provide measures of mental maturity and developmental level, and second, as projective devices for assessing personality characteristics. For measuring mental maturity and developmental level, the Goodenough Draw-A-Man Test (Goodenough, 1926; Goodenough and Harris, 1963) has been most frequently used. As a projective personality assessment, the Machover Draw-A-Person Test (Machover, 1949) has received widespread application. A recent outgrowth of the Draw-A-Person was the Koppitz procedure for scoring results from children ages five to twelve (Koppitz, 1968). Koppitz attempted to achieve a relatively simple and objective scoring system for personality and developmental assessment referred to as the Human Figure Drawing Test. The developmental score yielded by the Koppitz procedure is actually a shortened version of the Goodenough-Harris procedure.

Psychologists and educators have studied and used human figure drawing as assessments of personality and development for some time. Since the beginning of systematic attempts to study the usefulness of human figure drawings as a measurement device late in the 19th Century, there has been an "inexorable though slow increase in the consensual and empirical validity of projective drawings" (Ogden, 1976). Ogden, a recognized authority on psychodiagnostics, reported that many studies conclude human figure drawings are sufficiently reliable for behavioral predictions. But there remain many unresolved problems relating to their use as reliable and valid assessment tools.

While much of the research that has been reported on HFD tests suggests that they possess fair levels of reliability and validity (Guinan and Hurley, 1965; Landsman and Dillard, 1967), it is important to be aware that children's drawings of the human figure vary in quality and content because of a number of conditions and events extraneous to the ascribed purpose of the test. For example, an important source of variation in the quality and detail of figure drawings results from differences among children in motivation, a consideration which especially confuses the interpretation of test-retest differences (Schubert, 1969).

Quality and details in drawings of human figures are also influenced by the level of sensory-motor maturation of the child performing the drawing task (Whitmyre, 1953). It is because of this sensitivity to sensory-motor maturation factors, that educators have sometimes used HFD test results as indices of learning disabilities. But while sensory-motor deficits are sometimes linked to "learning disabilities," many children of prekindergarten and kindergarten level simply lack sufficient sensory-motor readiness

to allow them to perform with the precision fine motor coordinations required by the HFD task. It therefore becomes impossible, on the basis of this evidence alone, to distinguish between children who are mentally retarded, slow in maturation, and those who are beginning to show the discrepancy between potential and achievement which underlies the concept of "learning disability." Applications of HFD tests as screening devices for "learning disabilities" are therefore questionable.

Educators have sometimes adopted the use of figure drawing tests under the mistaken assumption that being devoid of linguistic content they are "culture fair." Clinicians who use these tests regularly, however, observe that children include elements in their drawings with greater frequency if they have acquired words to conceptualize these elements. Also, there is evidence that culture influences perceptions and that the social experiences of children are reflected in elements children select for emphasis in their human figure drawings (Machover, 1943). Therefore, it can be argued that scoring systems which are based on narrow samples, including only small numbers of children who are culturally different, may lead to inaccurate conclusions concerning the presence or absence of deficits. The extent to which socio-cultural biases prevail in the scoring of HFD tests has been minimally researched.

Although educators have used human figure drawings primarily as evidence of developmental differences and mental maturity, psychologists use these assessments as evidence concerning the existence of personality traits. While an extensive research literature exists on these applications, there is as yet no fully adequate objective scoring system for HFD (Buros, 1972). Interpreting and scoring human figure drawings for diagnosis of personality

structure can only be accomplished by clinicians with extensive education and experience. Teachers generally lack this capability; however, teachers can use the Koppitz (1968) scoring system for the Human Figure Drawing Test to derive a gross index of adjustment suitable for screening purposes. But even under the most carefully controlled circumstances, conclusions regarding the structure of personality or the quality of emotional life derived from an analysis of HFD's are subjective and dependent upon the skills, orientations, and possible biases of the scorer.

It is noteworthy that typically when HFD assessments are mentioned as a means of assessing personality or developmental level, a warning is issued that these indicators should only be considered part of a larger array of data. When used alone for diagnostic or classification purposes, HFD tests are misused. Taken alone, they fail to validate a diagnostic conclusion.

For assessing mental maturity an objective scoring system has been devised for the Goodenough Draw-A-Man test. Scoring a test requires about ten minutes of a teacher's time, assuming that the teacher has been instructed in the scoring system and has had an opportunity to score samples of student performance under supervised conditions. The interrater reliability of the Draw-A-Man scoring system is considered adequate for a gross screening, but the results produced by an individual child may be distorted by situational factors such as interest in the task, conditions under which the test is given, rapport with the examiner, fatigue, and perceived reinforcers. Perhaps as a consequence of these contaminations, a wide range of reliabilities has been reported, from approximately .20 to .70.

Research involving the use of HFD's as an index of general intelligence reveals that the strongest relationship occurs between these variables at ages six and seven. By age eleven, the consistency between these measures begins to fade. During adolescence, drawing tasks become poor indices of intellectual functioning because of changes in social roles. For older children, drawing skill becomes less salient. Sex differences begin to consistently appear, males exerting less effort in the production of adequate drawings (Hornowski, 1970).

Recent applications of HFD's in educational settings involve using the assessments to screen children into "risk categories" rather than screening for discrete learning impediments such as mental retardation, brain damage, or learning disability. The Evanston Early Identification Scale (Landsman and Dillard, 1967) has been offered as a "simple, highly efficient screening device for identifying those children who can be expected to have difficulty in school." The authors warn potential users that the scale is not a diagnostic instrument and that those performing poorly may do so for perceptual, social, or intellectual reasons. As a test for five and six-year-old children, the Evanston Early Identification Scale appeals to teachers because it can be administered to a group and because it has a clear scoring system with cut-off scores for the three categories: high risk, middle risk, and low risk.

Recommendations:

1. Human Figure Drawings can provide cost-efficient screening for five and six-year-olds if they are properly administered and scored, and if their use is limited to those tests with established validity. The fact that these measures have achieved wide popularity, as evidenced from the New York Survey, suggests that educators are convinced that they yield useful information. Their wide acceptance among educators argues for their retention in a screening battery if they are used under controlled conditions.

2. Because scoring systems differ, it is likely that different HFD tests will result in different conclusions regarding an individual child. Of course the present lack of comparability between school districts that use different HFD scoring systems makes for difficulty in interpreting variations in incidences of learning impediments.

3. If a single HFD procedure were to be selected, the Koppitz Human Figure Drawing Test would provide an acceptable option because: 1) it has received fairly wide application and acceptance; 2) it can be administered and objectively scored by teachers; and 3) it yields evidence of both intellectual and emotional learning impediments.

4. Teachers involved in screening should receive in-service instruction on administering and scoring whatever test is selected. It is unlikely that teachers will have acquired sufficient skill in the use of these instruments in their teacher preparation programs to enable them to conduct adequate screening.

5. Even though HFD tests contain no linguistic elements, teachers must be sensitive to the possibility that socio-cultural factors may influence test outcomes. Evidence regarding the presence of socio-cultural biases in these tests is limited and equivocal.

6. Under no conditions should evidence from human figure drawing tests alone be used for purposes of differential diganosis or placement. These tests must be used in combination with more detailed diagnostic testing to reveal the actual nature of the child's deficits.

7. The fact that these tests lack linguistic and auditory components make them unsuitable for use as the sole screening instrument for learning impediments. Overwhelming evidence pertaining to the structure of human abilities indicates that children who successfully pass screening with HFD tests may experience educational difficulties due to language impairments, auditory perception problems, and inability to deal with conceptual material, among other things.

[9]Tests used in New York State screening which involve human figure drawings are:
Draw-A-Man
Draw-A-Person
Metropolitan Readiness Test
Vane Kindergarten Test
ABC Inventory
Gesell Developmental Schedules
Denver Developmental Screening Test
First Grade Screening Test
Ilg and Ames
Riley Preschool Development Inventory

Ilg and Ames (Gesell Developmental Examination)[10]

The Ilg and Ames (Gesell Developmental Examination) is a clinical type of assessment designed to determine the general developmental level of the child. It is administered individually and requires the presence of a trained examiner in order to record and interpret the results. The primary use of the test is to detect whether a child is appropriately placed in a grade that is compatible with his developmental level. Ilg and Ames were very concerned with the damaging effects of overplacement i.e., grade placement above developmental level.

The Ilg and Ames includes the following performances areas:

1. The initial interview: Questions about age, birth date, birthday party including favorite activity and present received; siblings' names and ages; father's occupation.

2. Pencil and paper tests: Writing name or letters and address; numbers 1 to 20; copying six basic forms (circle, cross, square, triangle, divided rectangle, diamond in two orientations), and two three-dimensional forms (cylinder and cube in two orientations); completing Incomplete Man figure and giving his facial expression.

3. Right and left: (adaptation of Jacobson's Right and Left tests) naming parts and sides of body, carrying out single and double commands, responding to a series of pictures of a pair of hands in which two fingers are touching. Response is first verbal and then motor.

4. Form test: Visual One (Monroe) - matching forms; Visual Three (Monroe) - memory for designs; projection into forms.

5. Naming of animals for 60 seconds.

6. Concluding interview: Reporting on what child likes to do best in general, at school indoors and outdoors and at home indoors and outdoors.

7. Examination of teeth: Recording of both eruption and decay or fillings.

The above description of the test is taken from Frances L. Ilg and Louise Bates Ames, School Readiness: Behavior Test Used at the Gesell Institute. New York: Harper and Row, 1964. Test materials and forms are marketed by Harper and Row Publishers.

"It is an examination that is easier to give than to interpret, which is probably true of examinations in general. At one time we thought it might be possible to put this examination into the hands of the classroom teacher. Now we feel this should be done only under supervision."

The Ilg and Ames is of limited value in identifying specific areas of weakness or learning impediments. "It can be of great value when a child is having trouble in school, by giving information about his actual developmental level and identifying instances of gross misplacement. However, to be capable of conducting and evaluating a development test is one thing; to put it to practical use is quite another. As we have used these developmental tests in our clinical service we have often had difficulty in putting into practice the course of action which developmental findings might suggest."

[10]School Readiness (Gesell Developmental Examination), Frances Ilg and Louise Ames, Harper and Row, Inc.

Illinois Test of Psycholinguistic Ability[11]

In working to increase the rate of mental development of young disadvantaged mentally retarded children, the senior author of the Illinois Test of Psycholinguistic Ability noticed that many children who were labeled mentally retarded displayed wide discrepancies among their abilities. This was a quarter of a century ago, and although some of these discrepant abilities were spotted by informal methods, the need was felt for a systematic, diagnostic device which would tap and differentiate various facets of cognitive ability.

In 1961, the early form of the ITPA was developed and published in an experimental edition as a diagnostic test of communication abilities. The Revised Edition, as well as the original, was conceived as a diagnostic tool. Its objective is to delineate specific abilities and disabilities in children in order that remediation may be undertaken when needed. The ITPA provides 1. a framework within which tests of discrete and educationally significant abilities have been generated and 2. a base for the development of instructional programs for children.

The ITPA is used to delineate areas of difficulty in communication more than to determine overall ability. It is a diagnostic test of specific cognitive abilities as well as a molar test of intelligence.

The psycholinguistic model on which the ITPA is based attempts to relate those functions whereby the intention of one individual is transmitted (verbally or nonverbally) to another individual, and reciprocally, those functions whereby the environment or the intentions of another individual are received and interpreted. It deals with the psychological functions of the individual which operate in communication activities.

The present model of the ITPA postulates three dimensions of cognitive abilities.

1. Channels of Communication - the routes through which the content of communication flows. The channels may include various combinations of sensory input and response output. The major modes of input are auditory and visual; those of output are vocal and motor.

2. Psycholinguistic Processes - In analyzing behavior which occurs in the acquisition and use of language, three main processes are considered: a) the receptive process - that ability necessary to recognize and/or understand what is seen or heard; b) the expressive process - those skills necessary to express ideas or to respond either vocally or by gesture or movement; c) an organizing process which involves the internal manipulation of precepts, concepts, and linguistic symbols. It is a central mediating process elicited by the receptive process and preceding the expressive process.

3. Levels of Organization - The degree to which habits of communication are organized within the individual determines the level of functioning. Two levels are postulated in the clinical model of the ITPA; a) the representational level, which requires the more complex mediating process of utilizing symbols which carry the meaning of an object; b) the automatic level, in which the individual's habits of functioning are less voluntary but highly organized and integrated.

Functions tested by the ITPA:

I. Representational Level:

A. Receptive Process

1. Auditory Reception: Assesses the ability of a child to derive meaning from verbally presented material.

2. Visual Reception: Measures a child's ability to gain meaning from visual symbols.

B. The Organizing Process

1. Auditory-Vocal Association: Taps the child's ability to relate concepts presented orally.

2. Visual-Motor Association: Assesses the child's ability to relate concepts presented visually.

C. The Expressive Process

1. Verbal Expression: Assesses the ability of the child to express his own concepts vocally.

2. Manual Expression: Taps the child's ability to express ideas manually.

II. Automatic Level:

A. Closure

1. Grammatic Closure: Assesses the child's ability to make use of the redundancies of oral language in acquiring automatic hibits and handling syntax and grammatic inflections. a) Supplementary Test 1: Auditory Closure: Assesses the child's ability to fill in missing parts which were deleted in auditory presentation and to produce a complete word. b) Supplementary Test 2: Sound Blending: Assesses the organizing process at the automatic level. The sounds of a word are spoken singly at half-second intervals and the child is asked to tell what the word is.

2. Visual Closure: Assesses the child's ability to identify a common object from an incomplete visual presentation.

B. Sequential Memory

1. Auditory Sequential Memory: Assesses the child's ability to reproduce sequences of nonmeaningful figures from memory.

2. Visual Sequential Memory: Assesses the child's ability to reproduce sequences of nonmeaningful figures from memory. The ITPA is designed for individual administration to children from 2 - 10 years of age, and in the case of a more able child the easier items need not be given.

[11]Illinois Test of Psycholinguistic Ability, Kirk, McCarthy, Kirk, University of Illinois Press.

Meeting Street School Screening Test (MSSST)[12]

A short screening battery designed to identify children who are "high risk" in terms of potential learning disorders. An individual test that can be administered by teachers, the MSSST yields four scores:

1. Motor patterning
2. Visual-perceptual motor
3. Language
4. Total

The MSSST is a norm-referenced test standardized on 500 children (5-0 to 7-5 years of age). Fifty males and fifty females at each 6-month level were selected to represent the national population.

The test has been rated as lacking in suitable reliability and validity data. Of particular concern is the extent to which scores correlate with more psychometircally refined tests such as the WISC. Just what the test measures is unclear. No solid evidence has been collected showing the test's ability to predict future learning problems.

[12]Meeting Street School Screening Test, Peter K. Hainsworth & Marian L. Sigueland, Cripped Children & Adults of Rhode Island, Inc., Providence, Rhode Island.

Metropolitan Readiness Test[13]

The Metropolitan Readiness Test is one of the most widely used and more technically adequate measures of readiness for formal instruction in skill subjects such as reading. Composed of the sub-tests word meaning, listening, matching, alphabet, numbers, copying, and an optional figure drawing, the Metropolitan Readiness Test provides educators with a general index of expected performance in first-grade classroom situations. While the Metropolitan is composed of sub-tests, users are cautioned to beware of placing too much credence in sub-test scores themselves because of their limited reliabilities. For the test as a whole, however, the realiability has been judged adequate for its intended purpose.

The Metropolitan is designed as a predictive test, its main purpose being to differentiate between subjects of varying levels of expected performance. Because it predicts expected levels of later performance, the Metropolitan can serve as a screening device. It is important to recognize, however, that the Metropolitan is typically administered near the end of kindergarten. This procedure limits the test's usefulness for targeting children for diagnostic follow-up and early intervention programs.

Some socio-cultural distortions have been noted in test results, especially in lowered test reliabilities for children of disadvantaged socio-economic backgrounds. The test has been critisized for failing to yield information that would be of diagnostic value to teachers, even though the predictive validity of the test is considered good.

Of all individual published tests identified in New York State Survey, the Metropolitan Readiness Test ranked first in frequency of ap-

plication. Its position within the testing programs of school districts seems firmly established.

Because the content of the items within the Metropolitan is so similar to the kinds of tasks children are required to perform in the lower grades, the test is likely to detect incipient patterns of school failure. To specify the exact type of learning impediment responsible for inadequate performance on the Metropolitan would involve speculations beyond the intended use of the test. A content analysis of the test indicates that it would be sensitive to: 1) learning disabilities; 2) mental retardation; 3) sensory-motor integration; 4) visual and auditory perception; 5) language development; and 7) previous academic experience. In addition to these conditions and impediments, poor performance on the Metropolitan could result from: 1) bilingualism; 2) maturation lags; 3) inadequate teaching; and 4) narrow range of general experiences.

[13] Metropolitan Readiness Test, Gertrude Hildreth, Nellie Griffiths, and Mary McGauvran, Harcourt Brace Janovich, Inc.

Otis-Lennon Mental Ability Test[14]

The Otis-Lennon Mental Ability Test is a group test of high technical excellence designed to measure "verbal-education" mental abilities. All test items for kindergarten and grade one use pictorial stimuli. The teacher administers the test in a group setting and tasks are presented orally; no reading is involved.

The test is norm referenced and the standardization sample (Kdg. N=5,370; First grade N=710,000) is representative of the U. S. school population. Primary I battery dates 1967-69. Scores on the Otis-Lennon are expressed as: 1) deviation IQ's; 2) age percentile ranks; 3) grade percentile ranks; and 4) stanines.

Included in the test are items assessing.

1. Mental Processes of Classification
2. Following Directions
3. Quantitative Reasoning
4. Vocabulary
5. Semantic Relations

The test is administered with time limits, but research on the effect of speededness shows gains of only 2.9 IQ points when the test is administered untimed.

It is among the best tests of its type in reliability and validity (normed/technical excellence). Instructions for administration are clear and the test itself is well designed.

There is little evidence of socio-cultural biases in the test other than would result from differences in receptive language, particularly conceptual terms such as: center, least, next to, etc.

[14]Otis-Lennon Mental Ability Test, Arthur Otis and Roger Lennon, Harcourt Brace Jovanovich, Inc.

Peabody Picture Vocabulary Test[15]

The Peabody Picture Vocabulary Test (PPVT) estimates a child's verbal intelligence on the basis of his receptive vocabulary and his ability to match words that he hears with pictures that he sees. The test is considered to have special value for subjects with expressive language disorders and, since the child is not required to read anything, it is useful with non-readers. Because of the absence of expressive language, autistic, withdrawn, psychotic, and cerebral palsied persons are less handicapped in the testing situation. The subject must be able to hear words, see the drawings, and indicate "yes" or "no" in a manner which communicates.

The PPVT was standardized on 4,012 cases during 1958, using four examiners. Only white youth were included in the final standardization group. However, certain precautions were later taken to improve the generalizability of the norms.

Administering the PPVT requires no special preparation other than complete familiarity with the test materials, including practice in giving the instrument prior to its use in an assessment situation. Pronunciation is important, but if the instructions are closely followed the test can be given by a variety of professionals and paraprofessionals. Testing usually requires about 15 minutes since a basal/ceiling procedure is used. Less than the total test is therefore administered to any subject. Test administration must be done individually, however.

[15]Peabody Picture Vocabulary Test, Lloyd Dunn, American Guidance, Inc.

Perceptual Forms Test (Winter Haven)[16]

The Perceptual Forms Test (PFT) is for kindergarten, first grade or home administration and attempts to provide a measure of visual perception and school readiness. Individual or group administration of the test by teachers are possible. Testing and training are mixed together in the handbook resulting in a confused presentation.

Basically, the test involves copying geometric forms and completing incomplete forms. Responses are scored using a system related to the Koppitz: shape distinction, size distortion, rotation, and integration. No norms are provided with the test; instead the user is asked to divide the class into two groups, poor-to-average and average-to-excellent, based on the performance of the class as a whole. No convincing reliability or validity data is available on the test and it has been recommended that the Koppitz Gender-Gestald or the DTVP (Frostig) are more adequate tests serving the same purpose as the PFT.

[16]Perceptual Forms Test, Winter Haven Lions Research Foundation.

The Pupil Rating Scale[17]

The Pupil Rating Scale (PRS) for learning disabilities is a 24-item (8 sub-areas) "indirect" teacher rating scale designed to screen for learning disabilities and other learning impediments. Completing the scale requires approximately 10 minutes of a teacher's time for each student. Ratings can be conducted at leisure since judgements are made in subjective recollections of a student's performance. Although no information is provided on when to conduct the ratings, the scale is purportedly useful for normal children of ages four to eight.

The following scores are yielded by the scale:

1. Auditory comprehension
2. Spoken language
3. Orientation
4. Motor coordination
5. Personal-social behavior
6. Verbal subtotal
7. Nonverbal subtotal
8. Total

No major evaluations of the test are available beyond the report on the federally funded project that supported the development of the scale and an article in the Journal of Special Education "Test Review No. 15: The Pupil Rating Scale: Screening for Learning Disabilities."

Evidence reported in Final Report No. 1970-0-384-565 (DHEW), Minimal Brain Damage in Children relates to the validity of the scale and argues

that the scale compares favorable in validity with more elaborate, expensive, and time-consuming objective tests addressed to "fractionated" behavior. No acceptable evidence of the reliability of the rating is provided.

Only limited norms are available, and means and standard deviations for the five components are calculated on narrow groups. Further work in developing norms is necessary before the scale will fulfill its potential as a broad screening device.

The orientation of the test is toward real life functioning in the class-room which should lend appeal to teachers (in contrast to assessments based on obtuse theoretical considerations).

[17]The Pupil Rating Scale, Helmer R. Myklebust, Grune and Stratton.

Purdue Perceptual Motor Survey[18]

Wishing to avoid being held to standards of reliability and validity normally required of educational assessments, the author proclaims that the Purdue Perceptual Motor Survey (PPMS) is not a test. It is merely an instrument which permits the examiner to observe various behaviors that occur during structured activities. The "non-test", divided into eleven parts, consists altogether of twenty-two scorable items. The PPMS, generally used with second, third, and fourth graders, must be individually administered. As this survey purportedly diagnoses rather than screens, the author suggests that the user judge the value of administering the survey in terms of the child's specific needs.

Kephart's book, The Slow Learner in the Classroom, describes the test, underlying theory, and provides an analysis of poor achievement based on sensory motor patterns. Kephart focuses on the following major areas of the test: laterality, perceptual-motor matching, and directionality.

The PPMS, based on a quasi-scientific analysis of human learning is couched in oversimplified neurophysiological jargon. For example, it is postulated that perceptual organization develops in a child after the development of motor patterns.

Initial perceptual information is merely a "surge of energy" in the cortex's "projection areas." Perceptual data and motor data must match. Without this unity, the child exists in "two worlds" and thus, confusion occurs. Matching must begin with perception and unite with the motor. Since motor information is acquired first, the perceptual data is matched to it. This matching process is very detailed, involving large areas of the nervous system. The tactual and kinesthetic systems are integrated

first in the matching process. Visual and auditory systems, which have organized information, develop next. Because information concerning outside objects is readily available to the child, learning of form perception is possible.

Directionality is a learned behavior. It is developed from perceptual-motor match. Right-left, up-down, before-behind are the three coordinates of Euclidian space. The body develops these as a part of its motor patterns, then projects them on to outside space. This behavior is defined as perceptual projection, which is directionality. Objects acquire spatial dimensions through projection.

The following are some specific criticisms of the Purdue Perceptual Motor Survey:

1. The survey is specifically for the slow learner. It is not applicable to mentally retarded, physically handicapped, or blind children.
2. The survey is dominated with concern for laterality skills.
3. The Kraus-Weber items stress strength, not total physical fitness.
4. There is an age limit of six to ten, or second to fourth grades.
5. The scoring sheet needs revision. It is not convenient to score.
6. The manual is awkward to use while administering the test.
7. Some items can be repeated several times, which allows for the child to "think" of the wanted response.

This survey is not one that this writer would use as a regular classroom teacher. Most teachers are trained sufficiently to recognize, while teaching their normal curriculum, many problems which the survey could identify. For example, physical education would be an excellent place

to look for problems in maneuvering through an obstacle course, walking a balance beam, jumping, skipping, strength, imitation of movement (revealed in mimick games such as Simon says). The chalkboard exercises are assessable in writing skills, as in rhythmic writing. The ocular pursuit items would be taken by an optomcirist, since most teachers recommend eye tests immediately as failures occur. The Visual Achievement form consists of merely copying shapes in an organized pattern. A child does this several times a day, while copying from the chalkboard.

[18]Purdue Perceptual Motor Survey, Eugene Roach and Newell Kephart, Charles E. Merrill Co.

The Rutgers Drawing Test[19]

In its present form the Rutgers Drawing Test is considered a clinical tool rather than a screening instrument. A non-verbal test in which the S reproduces two-dimensional drawings (geometric) of increasing complexity, the Rutgers Drawing Test assesses:

1. Form Perception
2. Use of space
3. Motor coordination
4. Attention

Claims are made to the effect that the test can be administered in a group situation, but it basically remains an individual test. Form A is suited for children ages four to six.

As a drawing test (visuomotor) it is subject to some of the same problems of measurement as are human figure drawing tests (see previous discussion of HFD's) and taps similar aptitudes. The test may be useful in detecting evidence of emotional disorders where used by an experienced clinician, but basically reveals the S's visual percpetion and fine motor control.

No extensive reliability analysis has been done on the test, but Form A for young children shows stability coefficients consistent with other drawing tests. Concurrent validity has been investigated vis-a-vis the Stanford Binet and again shows results similar to human figure drawing tests.

Since the test is non-verbal, it may be suited for use with children of different linguistic or cultural backgrounds. The test's ability to serve as a means of differential diagnosis has not been established.

[19]The Rutgers Drawing Test, Anna Spiesman Starr, The Author.

Screening Tests for Identifying Children with Specific Language Disability (Slingerland)[20]

The Screening Tests for Identifying Children with Specific Language Disability, commonly referred to as "The Slingerland", is designed to screen for learning disabilities by targeting on a set of learning impediments presumed characteristic of the disorder: poor use of language coupled with inadequacies in writing, spelling, sensory-motor integration, and visual/motor patterning. Item content is somewhate redundant on the Illinois Test of Psycholinguistic Ability but the, Slingerland is a group test which makes it more suitable for screening applications. Although the test can be administered by teachers, potential users would be well advised to provide structured study for examiners prior to administration. Complexities in scoring, organization, and administration are numerous.

Eight sub-test are included in the battery:

1. Visual copying far point
2. Visual copying near point
3. Visual perception-memory
4. Visual discrimination
5. Kinesthetic memory
6. Auditory recall
7. Beginning and ending sounds
8. Auditory associations

Twenty-one different types of scores are generated for the above subtests, illustrating the level of complexity involved.

No reliability and validity data are reported on the test, along with an absence of norms. By proclaiming that the battery is not a "norm referenced instrument," the originator does not, however, avoid the need to know how much of the differences in performance that are noted among the children taking the test are due to error factors such as: presophistication, function fluctuation or unwanted traits, uncontrolled administrative conditions, scoring errors, etc. (unreliability). Neither is the lack of reliability and validity analyses compensated for by a "utility" analysis, showing how the results of testing can be put to profitable use in the classroom.

[20]**Screening Tests for Identifying Children With Specific Language Disability, Beth Slingerland, Educators Publishing Service, Inc.**

SRA Primary Mental Abilities (PMA)[21]

A test of specific intellectual abilities based on a multifactor theory of intelligence, the PMA measures (at grades K-1):

1. Verbal meaning
2. Perceptual speed
3. Number facility
4. Spatial relations

Reviewers of the test have been unimpressed with its failure to be improved over the last thirty years in response to detailed professional criticism. According to Quereshi, reviewing the test for Buros, "Since differential measurement by means of PMA, especially at the elementary school levels, is simply a will-o'-the-wisp, it is justified to use a general intelligence test which has more adequate standardization data and employs more defensible test construction procedures." And Shultz concludes in Buros, "further factor analytic work, exemplified by the programmatic efforts of Guilford and Cattell, showed clearly that the factors represented in PMA account for only fragmentary and arbitrary facets of cognitive space...consequently the PMA is currently analogous to the Ford Model A."

[21] SRA Primary Mental Abilities, 1962, Editor, L.L. Thurstone and Thelma Gurnin Thurstone, Science Research Associates, Inc.

Stanford Early School Achievement Test[22]

The Stanford Early School Achievement Test is a group achievement test suited for administration is grades kindergarten through 1.5. The child's cognitive activites can be measured on entry to kindergarten or entry to first grade. Heavy use of pictorial items is entailed in the test along with items which assess auditory skills, an area often ignored in early childhood screening instruments

Subtests are:

1. Environment (information items concerning social and natural science);
2. Mathematics (quantitative concepts);
3. Letters and Sounds;
4. Aural Comprehension (inference and sequence questions concerning aural passages).

Directions for administering the test are clear and children are provided with practice items. A few items have been criticized for social class biases.

Little validity information is provided with the test. Items reflect the content of kindergarten programs as interpreted by the test authors. High intercorrelations among the subtests seem to indicate the presence of a "g" factor within the battery and suggest the usefulness of the total score only for purposes of interpretation.

For the beginning of kindergarten, split-half reliabilities hover around the .80 level.

Suggestions for interpreting scores and planning instruction based on test performance are provided.

22 Stanford Early School Achievement Test, Richard Madden and Eric Gardner, Harcourt Brace Jovanovich, Inc.

Vane Kindergarten Test[23]

A partly group and partly individually administered test for children ages 4-0 to 6-11, the Vane Kindergarten Test measures:

1. Perceptual/motor skills (the child is required to copy a box, a cross, and a hexagon);
2. Vocabulary (eleven items);
3. Drawing a man (similar to the Goodenough Intelligence Test and the Goodenough-Harris Drawing Test);
4. Specific behaviors such as attention, ability to follow directions, self-control, cooperation, hand dominance, and speech are required to be noted by the examiner.

Results from the test can be converted to IQ scores which correlate approximately .75 with the Stanford Binet. However, the minimum IQ that four year olds can achieve is 100 which limits the test's usefulness for pre-kindergarten applications. The suitability of the test as a screening instrument is undermined by the fact that it is intended for use by school psychologists, not teachers. Time requirements for administration are short, however, the directions are simple and the screening rapid.

A limited number of studies have been made of the test's reliability and validity, reliability analysis being confined to stability (test-retest) with small groups. The primary means of validation has been "concurrent," against more elaborate tests of general intelligence. Both the reliability and validity of the test would be considered fair, with further work being needed in these areas.

The test was standardized on 400 children in New York and New Jersey with control for socio-economic status and ethnicity (white, non-white).

White/non-white distribution in the standardization group approximated closely the distribution prevailing in the 1960 U.S. Census. While rated by Buros, the Vane is not among those studied by UCLA's Center for Evaluation and Early Childhood Research Center. See also discussions of: Human Figure Drawing Tests and Peabody Picture Vocabulary Test.

[23]Vane Kindergarten Test (1968), Julia R. Vane, Clinical Psychology Publishing Co., Inc.

Wepman (The Auditory Discrimination Test)[24]

The Auditory Discrimination Test was developed by Joseph M. Wepman in 1958 and purports to identify children from five to eight years of age having auditory discrimination deficits. Each form of the test (1 & 2) contains 40 items comprising three-to-five letter word pairs. The words comprising each pair were matched for 1. familiarity, as determined by position in the Lorge-Thorndike list, 2. membership in the same phonetic category, and 3. Length. Vowel comparisons were based on parts of the tongue raised, height of tongue, and position of the lips.

The test is administered individually and requires from five to ten minutes to complete. A period of practice precedes the test. The child is instructed to listen to words spoken two at a time by the examiner and to indicate whether the words read were the same or different. If they were the same then the response should be positive (i.e. the child says "yes", "the same", or shakes his head in a positive manner). When the examiner thinks the child understands what is being asked of him, he proceeds by giving him some sample word pairs which are listed in the booklet if he discriminates correctly or demonstrates he understands the directions, then go on with the test. If the child discriminates incorrectly or shows in some way he is not comprehending, the tester reports the instructions, trying to simplify them.

The score is determined by the number of errors made distinguishing between words which are different and the same. Inadequate auditory discirmination is indicated when five year old children make more than six errors; six year olds make more than five errors; seven year olds make more than four errors; and children eight years old or older make more than three errors.

To be taken into consideration when contemplating using the Auditory Discrimination Test:

1. The test is poorly standardized
2. Inflection of the voice, such as pitch, volume, enunciation and other individual speech characteristics, may affect the performance of the subject.
3. The test meausres discrimination of words and not other auditory sounds.
4. There is no correlation between test and final hearing diagnosis by specialists.
5. The concepts of same and different would have to be understood.

Among the advantages ascribed to the test are:

1. Test-retest showed a reliability of .91.
2. The test does not require motor coordination or vision, speech, or reading ability.
3. The test is easy to administer. It takes only five to ten minutes and can be given in the classroom.
4. The test is inexpensive.

The Auditory Discrimination Test provides a quick assessment of auditory discrimination among children from five to eight years of age. Deficits in auditory discrimination have been found to be both related to the development of speech accuracy and somewhat related to reading ability. The information the author presents under validity also suggest relationships between auditory discrimination, intelligence and hearing.

[24]Auditory Discrimination Test, Joseph Wepman, Language Research Associates.

Wide Range Achievement Test[25]

The Wide Range Achievement Test (WRAT) is a teacher administered, part individual and part group test. Two levels of the test are available in the 1965 edition. Level I is for ages five to eleven, level II for ages twelve and above.

Administration procedures are complex because some parts are administered individually and some in a group situation under specified conditions. Not every testee responds to the same items, later item selection being dependent upon earlier performance. The instructions are confusing.

Items in the WRAT are:

1. Copying (pre-spelling)
2. Printing or writing name
3. Pre-reading, reading alphabet
4. Arithmetic - counting, reading numbers, showing, adding, subtracting.

Speed of response is an important component of the test score which tends to increase its reliability at the expense of its validity.

Norms from seven states are available to aid in interpreting performance, but their adequacy is questionable because of lack of information concerning the normative sample's characteristics.

Validity of the WRAT has been widely criticized, even at the content level. Many of the items, e.g. copying abstract symbols under timed conditions, are more commonly found in aptitude batteries rather than achievement tests.

[25] Wide Range Achievement Test, Justak, J. F. & S. W. Bijou, Guidance Associates.

BIBLIOGRAPHY

Albee, G. and Hamlin, R. An investigation of the reliability and validity of judgements of adjustment inferred from drawings. Journal of Clinical Psychology, 1949, 5, 389-392.

Buros, O.K. (Ed.) The Seventh Mental Measurements Yearbook, New Jersey: The Gryphon Press, 1972.

Buros, O.K. (Ed.) The Sixth Mental Measurements Yearbook, New Jersey: The Gryphon Press, 1965.

Carmerm, N. Invdividual and social factors in development of graphic ability. Journal of Psychology, 1938, 5, 165-184.

Frankenburg, W., Goldstein, A. & Camp, B. The revised Denver Developmental Screening Test: its accuracy as a screening instrument. Journal of Pediatrics, 1971, 79, 988-995.

Gardner, W. Learning and Behavior Characteristics of Exceptional Children and Youth. Boston: Allyn and Bacon, 1977.

Goodenough, F. Measurement of Intelligence by Drawings. New York: Harcourt, Boace & World, 1926.

Goodenough, F. and Harris, D. Goodenough-Harris Dawing Test. New York: Harcourt, Brace & World, 1963.

Graubard, P. Children with behavioral disabilities. In L.M. Dunn (Ed.) Exceptional Children in the Schools: Special Education in Transition (2nd ed.). New York: Holt, Rinehart & Winston, 1973.

Guinan, J. and Jurley, J. An investigation on the reliability of human figure drawings. Journal of Projective Techniques, 1965, 28, 300-304.

Hill, J. Models for screening. Minnesota University, Minneapolis: Institute of Child Development, 1970. (ERIC Document Repruction Service No. ED 039 660)

Hoepfner, R., Stern, Carolyn & Nummedal, S. (Eds.). CSE-ECRC Preschool/Kindergarten Test Evaluations. Los Angeles: UCLA Graduate School of Education, 1973

Hornowski, B. Studies on Psychological Development of Children and Youth by Using Children's Drawings of a Human Figure. Warsaw: Polskiej Akademic Nauk, 1970.

Ilg, F. and Ames, L. School Readiness: Behavior Tests Used at the Gesell Institute. New York: Harper and Row, 1964.

Keogh, B. and Becker, L. Early detection of learning problems: questions, cautions, and guidelines. Exceptional Children, 1973, 39, 5-11.

Kephart, N. The Slow Learner in the Classroom. Columbus: C. E. Merrill Books, 1960.

Koppitz, E. Psychological Evaluation of Children's Human Figure Drawings. New York: Grune and Stratton, 1968.

Landsman, M. and Dillard, H. Evanston Early Identification Scale. Follett Education Corporation, 1967.

Machover, K. Personality Projection in the Drawing of the Human Figure. Springfield, Ill.: Charles C. Thomas, 1949.

Machover, S. Cultural and Racial Variations in Patterns of Intellect. New York: Contributions to Education, No. 875, 1943.

Myklebust, H. and others. Minimal brain damage in children, final report. Health, Education, and Welfare Dept.: Division of Chronic Disease Programs, Neurological and Sensory Disease, 1970.

Ogdon, D. Psychodiagnostics and Personality Assessment: A Handbook (2nd Ed.) Los Angeles: Western Psycholigical Services, 1976.

Proger, Barton B. Test review no. 15: the Pupil Rating Scale: Screening for Learning Disabilities, Journal of Special Education, 1973, 7, 311-317.

Reznikoff, M. and Toblen, D. The use of human figure drawings in the diagnosis of organic pathology. Journal of Consulting Psychology, 1956, 20, 467-470.

Schubert, D. Decrease in rated adjustment on repeat DAP tests apparently due to lower motivation. Journal of Projective Techniques and Personality Assessment, 1969, 33-34.

State University of New York. A Survey of Educational Screening Programs in New York State School Districts and Boards of Cooperative Services. Albany: The State Education Department, 1975.

Tarjan, G. In J. H. Meier (Ed.) Screening and assessment of young children at developmental risk. (The President's Committee on Mental Retardation, DHEW Publication No. (OS) 73-90). Washington D. C.: Government Printing Office, 1973.

Uhitmyre, J. The significance of artistic excellence in the judgement and of adjustment inferred from human figure drawings. Journal of Counsulting Psychology, 1953, 17, 421-424.

Vane, J. and Eisen, V. The Goodenough Draw-A-Man Test and signs of maladjustment in kindergarten children. Journal of Clinical Psychology, 1962, 18, 276-279.